Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI)
Vol. 9, No. 4, December 2023, pp. 941-950
ISSN: 2338-3070, DOI: 10.26555/jiteki.v9i4.26659

New Generation Indonesian Endemic Cattle Classification: MobileNetV2 and ResNet50

Ahmad Fikri, Aniati Murni
Universitas Indonesia, Pondok Cina, Beji, Kota Depok, Jawa Barat, 16424, Indonesia

ARTICLE INFO

Article history:

Received July 11, 2023
Revised September 25, 2023
Published October 10, 2023

Keywords:

Indonesian Endemic Cattle;
MobileNetV2;
ResNet50

ABSTRACT

Cattle are an essential source of animal food globally, and each country possesses unique endemic cattle races. However, categorizing cattle, especially in countries like Indonesia with a large cattle population, presents challenges due to costs and subjectivity when using human experts. This research utilizes Computer Vision (CV) for image data classification to address this urgent need for automatic categorization. The main objective of this study is to develop a mobile-friendly model using CV techniques that can accurately detect and classify Indonesian endemic cattle races, such as Limosin, Madura, Pegon, and Simental. To achieve this, an object localization approach is employed to extract multiple features from distinct regions of each cattle image, including the head, ear, horn, and muzzle areas. The authors evaluate two CV algorithms, ResNet50 and MobileNetV2, to assess their performance in cattle race classification. The dataset used is facial photos of 147 cows. The results indicate that ResNet50 outperforms MobileNetV2, achieving a training data accuracy of 83.33% compared to MobileNetV2's 77.08%. Moreover, the validation accuracy of ResNet50 (76.92%) significantly surpasses MobileNetV2's (38.46%). The novel contribution of this research lies in developing a cost-effective and efficient solution for identifying endemic cattle breeds in Indonesia. The mobile-friendly model based on ResNet50 demonstrates superior accuracy, enabling cattle farmers and researchers to categorize cattle races with higher precision, reducing manual effort, and minimizing costs. In conclusion, this research provides a valuable advancement in automatic cattle categorization using CV techniques. By offering a practical and accurate model that considers Indonesia's specific cattle breeding conditions, this study contributes to the sustainable management and conservation of endemic cattle races while enhancing the efficiency of cattle farming practices.

Corresponding Author:

Ahmad Fikri, Universitas Indonesia, Pondok Cina, Beji, Kota Depok, Jawa Barat, 16424, Indonesia
Email: ashterigaz@gmail.com

1. INTRODUCTION

Cattle play a pivotal role in global food consumption, producing meat and milk, making them vital export commodities in numerous countries. Each country harbors unique endemic cattle breeds, necessitating distinct cattle management approaches. In industrial settings, farms manage large cattle populations concurrently to ensure effective and resource-efficient processes. Cows are one of the prima donna animals for consumption globally [1]. Besides producing meat like chickens and goats, cows can also produce milk. As we know, cow derivative products are essential export commodities in several countries [2]. This fact makes the demand for cattle in various countries always urgent. Uniquely, cows in each country have endemic types. This results in each country having a unique need for cattle.

With Indonesia being the fourth most populous country globally, its growing cattle population requires innovative management methods [2]. Data collection is crucial in optimizing industry practices and generating new knowledge for organizations. Accurate cattle classification is among the challenges in cattle management, which aids in providing appropriate treatment. However, human-based classification can be prone to errors due to daily changes in cattle appearance and issues with tagging methods.

In this modern era, cattle management is attempted using innovative methods. Data collection is essential for the industry [4]. Another benefit is that the data collected will become new knowledge for the organization [5]. One of the big problems in cattle management is the classification of cattle [6]. The type of cattle is advantageous in providing proper treatment for the cow. It's just that management using humans is prone to mistakes because cows grow and develop daily, so their appearance changes [7]. When using ear tags, these objects are prone to falling off, so the cow cannot be recognized [8]. Body painting or stamping with a hot iron is also prone to loss due to cows' daily growth and development [9]. These three things are less effective for livestock conditions.

Notably, the research unveils the uniqueness of cattle noses, similar to fingerprints, enabling individual identification and recognition of changes as they develop. Leveraging this discovery, the authors propose processing cattle nose images through Computer Vision (CV) technology, particularly for automatic classification. The writer finds the uniqueness that the cow's nose is identical to the cow's. According to [3], [9], [10], the cow's nose pattern is unique, like fingerprints. That way, cows can be distinguished based on their noses, and they can still be recognized when they develop and develop. Therefore, it is necessary to process the image of each photo of the cow's nose which will later be stored in the database. With the help of this technology, also known as Computer Vision (CV), it is hoped that farms will increase profitability by eliminating operational costs.

The existing related works, such as those using Expectation Maximization (EM), Convolutional Neural Network (CNN), and Content-Based Image Retrieval (CBIR) approaches, have shown promising results in cattle classification. However, a crucial gap remains in using mobile-friendly CV models like MobileNetV2 and ResNet50 for cattle classification in the context of Indonesian endemic cattle breeds. The result of [11] shows the application of deep learning, especially CV, in treating livestock with two phases. These phases are the segmentation phase and the classification phase of feature extraction. The combination of these phases is to create more specific features by classifying livestock first. Hamdi and his team used Expectation Maximization (EM) for the segmentation algorithm and extracted texture features from each image. Furthermore, they use FKNN for the classification algorithm. This study's dataset consisted of 32 groups of cow snout images. FKNN achieves a maximum accuracy of 100%, higher than KNN with 88% accuracy.

Paper [12] used the technique as a feature extraction algorithm and a Convolutional Neural Network (CNN) to classify Indonesian endemic cattle. GLCM can extract these features: energy contrast and homogeneity via CNN. This study categorizes five types of cattle: Bali, Pasuruan, Aceh, Madura, and Pesisir. The experimental results show that combining GLCM and CNN has higher accuracy than the original CNN. The precise accuracy of the GLCM-CNN is 98.927% for 100 iterations of the image data set. However, [11] only performed feature segmentation after classifying. As a result, the model obtained is more sensitive to available features than specific features, such as differences in face shape and nose shape.

Research [13] focused on classifying cattle using the Content-Based Image Retrieval (CBIR) approach. In addition to the technical differences between [11] and [12], several steps were carried out in this phase to obtain classification work, such as preprocessing, significantly changing the size and color of the frame, and transforming the color space. Furthermore, Sutojo uses GLCM for feature extraction to find contrast, energy, homogeneity, and entropy. In this paper, Sutojo uses a sample of combined cattle species between endemic cattle (Limosin, Simental, Brangus) and cross breeds (Ongole (PO) and Frisien Holstein (FH)). Each cow has 100 training images and 20 test images. The results of this image processing are extraordinary, with 95% system accuracy and 100% precision.

Research on Angus Cettle [3] also strengthens the claim for face detection in cattle. In this study, the detection performed on Angus cattle obtained a better prediction accuracy on the face and better in combination with the body. With three types of algorithms, PrimNet, VGG16, and ResNet50, this algorithm achieves significant accuracy with the VGG16 algorithm.

However, the drawback of the [3], [11], [12], [13] models is that development is done using applications that are on the computer. Meanwhile, the required device characteristics are mobile devices capable of processing images with high flexibility and mobility, but the resulting accuracy is guaranteed [14], [15]. Two mobile algorithms can accommodate this need, namely MobileNetV2 and ResNet50. MobileNetV2 was proposed as a model by the Google Research team to effectively maximize accuracy while considering the limited resources for on-device or embedded applications [16]. Meanwhile, ResNet50 is unique because it already has an image dataset that is pre-trained and ready to use. Both algorithms are worth considering as future algorithms in classification [17]. Specifically, there is no use of the MobileNet or ResNet algorithm types in the classification of Indonesian cattle, although that is very urgent when you see the advantages. Even so, this algorithm has been used for a variety of human faces ([18]–[22]), fish ([23]), breast cancer ([24]), and human skin ([25]).

Thus, this research addresses this gap by developing a novel mobile-friendly CV model for cattle classification. The study focuses on four Indonesian endemic cattle breeds: Limousin, Madura, Pegon, and Simental. By utilizing MobileNetV2 and ResNet50, mobile-friendly deep learning models, the authors seek to enhance cattle breeders' profitability and efficiency in cattle management.

The authors contribute to advancing automatic cattle classification using state-of-the-art CV techniques through this research, offering a flexible and accurate solution for cattle management, promising cost-effectiveness, and increased efficiency for cattle farmers. For the case used, the authors took samples from four Indonesian endemic cattle races: Limousin, Madura, Pegon, and Simental. Meanwhile, there has yet to be similar research that addresses the same issue as in this study ([16], [18]–[29], [3], [8], [10], [16]–[18], [25], [26], [29]–[42]). The subsequent sections detail the methodology, experimental results, and discussions, providing comprehensive insights into the effectiveness of the proposed model.

2. PROPOSED METHOD

In this study, two convolutional neural network models, ResNet50 and MobileNetV2, classify cattle images based on their race. The ResNet50 architecture, known for its effectiveness in image recognition, uses residual learning to address the degradation problem in deeper networks. On the other hand, MobileNetV2, designed for mobile and embedded devices, utilizes depthwise separable convolutions and inverted residual blocks to reduce computational complexity and memory requirements while maintaining performance.

ResNet50 is one of the most popular and effective convolutional neural network model architectures for image recognition, especially on mobile devices [42]. This architecture was developed by Microsoft Research in 2015 and won the ImageNet image recognition competition the same year. The ResNet50 algorithm is designed with a residual learning form in the form of several stacked layers. Formally, this learning formulation is written in (1).

$$y = F(x, \{W_i\}) + x \tag{1}$$

From (1), x and y are the input and output vectors of the layer under consideration. The function $F(x, \{W_i\}) + x$ represents the residual mapping to be studied. For the example in Fig. 1 which has multiple layers, $F = W_2\sigma(W_1 x)$ where σ denotes ReLU [17] and the bias is omitted for notational simplification. $F + x$ operations are performed by shortcut connection and addition by element. We adopt the second nonlinearity after addition (that is, $\sigma(y)$).

To discuss residual learning, let's think of $H(x)$ as a base mapping to fit some stacked layer (not necessarily the entire net), where x denotes the input to this first layer. If one hypothesizes that some nonlinear layers can asymptotically approximate a complicated function[2], then that is equivalent to hypothesizing that they can asymptotically approximate the residual function, i.e. $H(x) - x$ (assuming that the inputs and outputs have the same dimensions). So instead of expecting nested layers to approximate $H(x)$, we explicitly let these layers approximate the residual function $F(x) := H(x) - x$. The original function becomes $F(x) + x$. Although both forms should be able to asymptotically approximate the desired function (as hypothesized), the ease of learning may differ.

Furthermore, the shortcuts in (1) introduce no additional parameters or computational complexity. These results are interesting in practice and important in computational comparisons between ordinary and residual networks. Authors can fairly compare regular/residual networks that simultaneously have the same number of parameters, depth, width, and computational cost (except for the addition of negligible elements). The x and F dimensions must be the same in (1). If this is not the case (e.g. when changing input/output channels), the author can perform a linear W_s projection with a shortcut connection to fit the dimensions:

$$y = F(x, \{W_i\}) + W_s x \tag{2}$$

In (1) calculations, the square matrix W_s can also be utilized. However, the authors will show through experimentation that identity mapping is sufficient to overcome the degradation problem and is economical, and thus Ws is only used when matching dimensions in (2). The form of the residual function F is flexible. The experiments in this paper involve function F having two or three layers (Fig. 1), while more layers are possible. But if F has only one layer, equation (1) is similar to a linear layer: $y = W_i x + x$, whose advantage we have not observed yet. The authors also note that while the notation above is about fully connected layers for simplicity, it applies to convolutional layers. The function $F(x, \{W_i\})$ can represent many convolutional layers. Addition by element is performed on two feature maps, channel by channel.

ResNet50 has 50 convolution layers used to process images and extract important features. This architecture uses the concept of a "skip connection" or "shortcut connection" (as illustrated in Fig. 1), which allows information to pass through multiple layers as it travels through the network [32]. This helps overcome the degradation problem in deeper convolutional neural networks.

In addition, ResNet50 also uses a convolution block called a "bottleneck," which consists of a convolution layer with a smaller filter followed by a convolution layer with a larger filter. These blocks help reduce computational complexity and speed up training time [17].

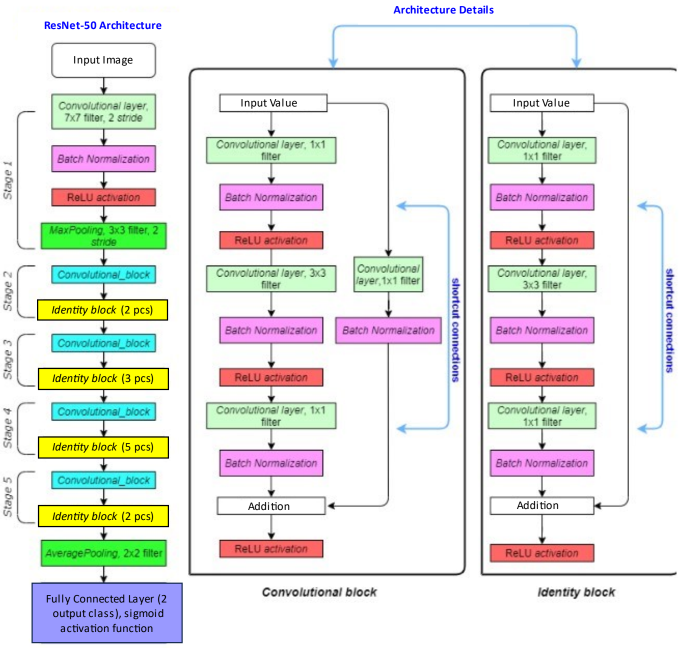

Fig. 1. ResNet50 Algorithm Scheme [30]

The following algorithm is MobileNetV2. This algorithm is designed as one of the architectural convolution neural network models for mobile devices or embedded devices with computational and memory limitations. This architecture was developed by Google in 2018 as an improvement over the previous MobileNet model [26].

MobileNetV2 uses a similar bottleneck layer concept to ResNet50 but with a smaller filter size to reduce the number of parameters and computational complexity [29]. This architecture also utilizes the "depthwise separable convolution" technique, which processes each channel in the image separately, thereby reducing the number of parameters and computation time. In addition, MobileNetV2 also introduces the concept of an "inverted residual block," which allows information to jump through multiple layers as it passes through the network, thereby reducing performance degradation in deeper networks [16].

Fig. 2 depicts the layer arrangement in MobileNet. Just like MobilenetV1, MobileNetV2 still uses depthwise and pointwise convolution [19], [21]. MobileNetV2 adds two new features, namely: 1) linear bottleneck, and 2) shortcut connections between bottlenecks. In the bottleneck section there are inputs and outputs between the model, while the inner layer encapsulates the model's ability to convert input from lower level concepts (for example pixels) to higher level descriptors (for example image categories). Ultimately, just like residual connections in traditional CNNs, shortcuts between bottlenecks allow faster training and better accuracy.

The modeling flow of the proposed system is depicted in Fig. 3. It involves several key steps, including data preprocessing, object localization, and feature extraction, followed by training and evaluation using the ResNet50 and MobileNetV2 models.

Fig. 3 illustrates the design of the proposed system to classify cattle images based on their race. In the first step, the image as input is entered to be processed. The shot entered is a picture of a cow's head, as shown in Fig. 4. Furthermore, the idea is mapped with the object localization process. This process is carried out to capture the crucial parts that will be processed into the algorithm. As mentioned, the image to be taken is the natural area, especially the nose, which must be visible so that the uniqueness of the cow can be optimally processed. Next, the third step is storing features ready to be processed. These features will be entered into a database ready to be separated into training, testing, and validation datasets. Finally, the feature will be used by two algorithms, namely the ResNet50 and MobileNetV2 algorithms. Both algorithms will be trained to classify cattle images based on race. The hope is that the outputs of the two models are then compared to determine the final classification of the input image.

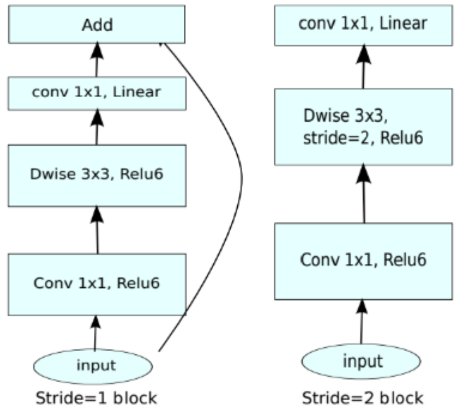

Fig. 2. MobileNetV2 Arrchitecture [24]

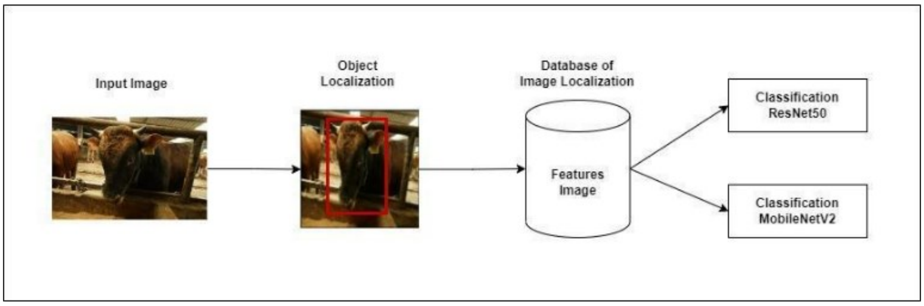

Fig. 3. Model Scheme

Fig. 4. Image Dataset of Limousin Cattle (a), Madura (b), Pegon (c), and Simental (d)

The author uses the Roboflow framework to process all data. Roboflow is an image annotation platform that facilitates image annotation tasks in developing computer vision models. Roboflow allows users to import image datasets, assign image annotation tasks, and generate pre-processed output. This study used a dataset of cow images annotated with labels to mark areas as local objects. The Roboflow framework is used to provide such annotations. Localized object areas correspond to the features to be extracted, such as head, hair, nose, eye, and horn areas. Once the local object is generated, we perform a pre-processing stage where we crop the image in the local area to prepare it as input for the testing process. The goal is to increase the system's accuracy by focusing on local points and eliminating unnecessary regions. The steps in more detail in this research are:

1. Datasets
   The study utilizes 147 images of cattle representing four Indonesian endemic cattle races: Limousin, Madura, Pegon, and Simental. The dataset is split into training, validation, and test sets in a ratio of 60:30:10, respectively.
2. Pre-processing
   In the data preprocessing stage, the cow images are resized to a uniform size of 416×416 pixels, ensuring consistency for object localization. The object localization process is crucial for identifying specific areas in the cow images, such as the head, hair, nose, eye, and horn areas, to extract relevant features. The authors use the Roboflow framework to achieve this, which allows image annotation tasks to mark localized object areas.
3. Localization of Objects
   The information in image data is very complicated because it requires generalizations to describe objects and pinpoint their exact location. To overcome this problem, the object localization method is used, which consists of three components that aim to improve accuracy: area of interest techniques that focus on specific areas to extract features, classification techniques, and object localization.

   In this process, each image data of the type of cattle will be labeled (annotated image) to determine the location of the features to be extracted from each image of the cow. Fig. 5 shows an example of an annotated Pegon cow image for a localization object. These object annotations identify specific areas of the image containing the thing or object the system wants to recognize. In this image, the cow object has been annotated to help the system identify the cow's location in the image. The second image (right) shows an example of a cropping image of a Pegon cow focusing on the localization area. After the object localization, the images are cropped to focus on the specific regions of interest, eliminating unnecessary regions and enhancing the system's accuracy.

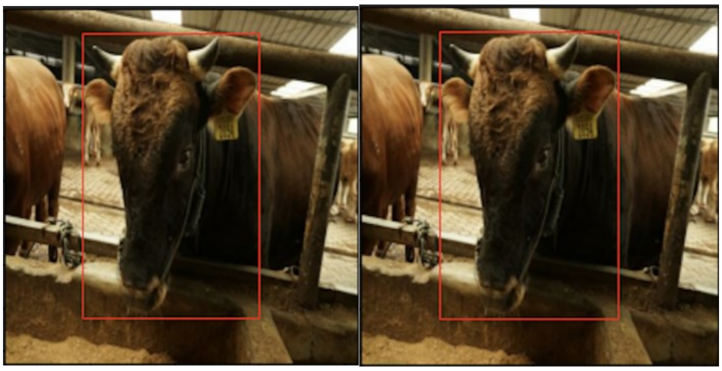

Fig. 5. Example of annotated cow image data for localization (left) and cropped (right) objects

4. Evaluation Score
   To evaluate the performance of the models, the accuracy metric is calculated (as shown in Equation 3), which measures the proportion of correctly classified cattle images to the total number of predictions. However, it is recommended to include other evaluation metrics such as precision, recall, and F1 score to provide a more comprehensive assessment of the models' performance.

$$Accuracy = (TN + TP) / (TP + FP + TN + FN) \tag{3}$$

In summary, the method section outlines using ResNet50 and MobileNetV2 models, the modeling flow, data preprocessing, object localization, and evaluation. However, to enhance clarity, provide more details on the object localization technique, training process, data augmentation (if applicable), and additional evaluation metrics used in the study.

3. RESULTS AND DISCUSSION

This experiment tests two parameters: the number of epoch iterations and the computing time. The number of epoch iterations shows how often the entire dataset is used to train the model, while the computation time shows how long it takes the system to process the data to produce the model. These two parameters are parameters that are commonly used in similar research, especially those related to biometric analysis [20], [21], [24], [25], [27].

The performance results from the system experiments show that the higher the number of epoch iterations, the better the model performance will be. In this study, the authors set the same epoch for both models, namely 400 n, as the default number for those models.

Regarding computation time, the shorter it takes to train the model, the more efficient the system is. Computational time efficiency does not always correlate with the resulting model performance. The short time may compromise the model's accuracy because it does not provide sufficient opportunities for the model to learn the complex data features.

Table 1 shows that in cases with the same number of epochs of 400, the MobileNet V2 model has a shorter computation time than the ResNet-50 model. The MobileNet V2 model requires approximately 35 minutes of computation time, while the ResNet-50 model requires about 120 minutes. This shows that the MobileNet V2 model is more efficient regarding computation time because it is specifically designed for mobile devices with limited computing resources. Meanwhile, the ResNet50 classification algorithm performs in-depth tests on all network layers.

Table 1. Comparison of Model Performance

Algorithm	Computation Time (m)
ResNet50	120 m
MobileNet-V2	35 m

Based on Table 2, ResNet50 is superior in terms of validation data accuracy. Even though both have accuracy above 70%, as is the standard accuracy of similar models, ResNet50 is far behind MobileNetV2 in calculating validation data accuracy. However, there is a uniqueness in the validation data figures for each model. ResNet50 has a validation accuracy that is close to the accuracy value on the testing data, while on MobileNetV2, the validation value is only half the accuracy value on the testing data. This shows that there is potential for overfitting of the model.

Table 2. Classification Model Accuracy Results

Algorithm	Accuracy (%)	Validation Accuracy (%)
ResNet50	83.33	76.92
MobileNetV2	77.08	38.46

Previous studies using objects similar to this research indicate that the potential for overfitting is close to the lack of datasets. However, if you look at it from another side, there is a potential method that can be used other than looking at the dataset, namely the use of regularization techniques. With regularization techniques, researchers focus more on finding patterns from the model rather than increasing the quantity of datasets which cannot necessarily be proven to be related to increasing model quality. The best way to reduce overfitting or the best way to set a fixed sized model is to get the average prediction of all possible parameter settings and combine the final output. However, this becomes too computationally expensive and not feasible for real-time inference/prediction. There are other methods inspired by ensemble techniques (such as AdaBoost, XGBoost, and Random Forest) where researchers use several neural networks with different architectures. But this requires many models to be trained and stored, which over time becomes a big challenge as networks grow deeper. Obviously, this conflicts with the concept of MobileNetV2 which offers a more computationally efficient model.

The use of MobileNetv2 and ResNet50 using regularization techniques has been discussed in several previous studies. One of the keys to the two models is that the dataset must be large because it learns from a large number of models [29], [30], [31]. In this research, limitations in collecting datasets are an issue due to the limited number of cows in each farm, even though this dataset has been collected from one of the largest farms in Indonesia. In neural networks, this is not an efficient use of hardware because the same features need to be created separately by different models. That's when the idea of using the same subset of neurons was discovered [41].

4. DISCUSSION

After developing and testing the model, the best model for this case study is divided into two: the model with the fastest processor and the model with the highest accuracy. For the model with the quickest process, MobileNetV2 takes the first spot. This is because ResNet50 has a much higher number of layers than MobileNetV2. We have seen from the different test results 70 minutes. This more than twofold difference indicates that model development with ResNet50 needs to be refined to catch up with MobileNetV2 gains.

On the other hand, the model with the highest accuracy is obtained by ResNet50. The more layers answer the reason for this gain [43]. ResNet50 is slower than MobileNetV2 in this case, but it has better accuracy on the accuracy and accuracy of validation data. Even though it has passed the 70% threshold [41], the ResNet50 model must be further developed because the validation accuracy is still lower than the training accuracy. On the other hand, MobileNetV2 must be set in the model structure because there is a very significant difference between training and validation accuracy. These improvements can be made by adding data, repairing the model, or adding a new layer to the model [44].

The comparison between the MobileNet V2 and ResNet50 models is inseparable from other factors, such as the accuracy and validation accuracy of the system test scenarios. Because one of the keys to the success of a cow image recognition system is the accuracy and validation accuracy of the system's experimental design on the cow image dataset used, the accuracy and validation accuracy values of the two models need to be considered and compared.

Please note that both models use machine learning, classified as deep learning. ResNet50 and MobileNetV2 use iterative learning up to the specified epoch point. The 400 epoch point was chosen because it is the default iteration value of the two models ([31], [34], [16]). One of the proper improvements to be made in future research is to determine the maximum iteration value for these two models.

The next development agenda is developing a model that can accommodate the two advantages of the model, namely the most accurate and fastest model. Considering the location of the case study, it is crucial to make the two go hand in hand. One of the subsequent research agendas is to develop a model to achieve both of these and improve the existing model. Not only that, but this model is also worthy of investigation with other case studies, such as other ruminant animal farms (goats, buffaloes, or sheep) or other similar case studies.

5. CONCLUSION

From the results of research conducted by the author, it can be concluded that the Computer Vision approach to the image dataset of cattle species with the localization object approach to Indonesian endemic cattle has two results. The most accurate model was achieved by the ResNet50 algorithm with a training data accuracy of 83.33% which outperformed the MobileNetV2 model by 77.08%. The validation accuracy value on ResNet50 was 76.92%, far superior to MobileNetV2 of 38.46%. On the other hand, the model with the fastest process was achieved by MobileNetV2 with 50 minutes compared to ResNet50's 120 minutes. This can happen because of the difference in the number of layers of ResNet, which is superior to MobileNetV2, which raises advantages and disadvantages simultaneously.

In general, the author succeeded in achieving the aim of this research, namely creating an automatic cattle classification using advanced CV techniques. The use of this algorithm in animal husbandry, especially the accurate classification of cattle based on race, can be beneficial for cattle breeders, animal husbandry and the agricultural industry in Indonesia because it does not injure the cattle and the mapped identity of the cattle is very unique and impossible to change. When used on a large scale, farmers can simply position one camera in the corral corridor and let the cows pass in the same direction from there. That way, the costs incurred for manpower costs can be eliminated.

This model can still be developed further to increase its accuracy and strengthen its specialization. This research encourages biometric identification of animals, especially cows, where cows are the largest and most popular livestock in the world after chickens and have different breeds depending on the continent. With a more general model like this, breeders are only limited to breeding cattle with the criteria for endemic Indonesian breeds. There are several ways to improve this model, such as using ensemble learning [25] or AutoML [26] which is also used in several research related to computer vision. As an extension of the application of computer vision which was initially used to identify human identity [27], research related to animal biometric identity is certainly very interesting.

The author suggests that further research focus on model development to improve accuracy and speed on both sides. From the validation accuracy rate, which is lower than the training accuracy, many things can be improved, such as the amount of data, number of layers, or model structure. In addition, this model is also feasible to be developed for other case studies, such as ruminant animals with identical characteristics similar to cattle (goats, bulls, or sheep).

REFERENCES

[1] A. Ma'arif, A. I. Cahyadi, S. Herdjunanto and O. Wahyunggoro, "Tracking Control of High Order Input Reference Using Integrals State Feedback and Coefficient Diagram Method Tuning," *IEEE Access*, vol. 8, pp. 182731-182741, 2020, https://doi.org/10.1109/ACCESS.2020.3029115.

[2] B. Basyar, "Beef cattle farm development policies to overcome beef distribution problem in indonesia: A literature review," *Am. J. Anim. Vet. Sci.*, vol. 16, no. 1, pp. 71–76, 2021, https://doi.org/10.3844/AJAVSP.2021.71.76.

[3] A. Agus and T. S. M. Widi, "Current situation and future prospects for beef cattle production in Indonesia - A review," *Asian-Australasian J. Anim. Sci.*, vol. 31, no. 7, pp. 976–983, 2018, https://doi.org/10.5713/ajas.18.0233.
[4] S. Chen, S. Wang, X. Zuo, and R. Yang, "Angus cattle recognition using deep learning," in *5th International Conference on Pattern Recognition (ICPR)*, pp. 4169–4175, 2021, https://doi.org/10.1109/ICPR48806.2021.9412073.
[5] J. Gabriel, A. Mayzira, J. Aditya, M. Itsari, S. Satrio, and Y. Ruldeviyani, "Critical Success Factors of Data Integration on Digital Human Capital Information System to Support Digital Transformation - A Case Study at PTXYZ," in *8th International Conference on Cyber and IT Service Management, CITSM*, pp. 1-7, 2020, https://doi.org/10.1109/CITSM50537.2020.9268793.
[6] F. Fauziyah, Z. Wang, and G. Joy, "Knowledge Management Strategy for Handling Cyber Attacks in E-Commerce with Computer Security Incident Response Team (CSIRT)," *J. Inf. Secur.*, vol. 13, no. 4, pp. 294–311, 2022, https://doi.org/10.4236/jis.2022.134016.
[7] M. Leach, S. Barney, S. Dlay, A. Crowe, and I. Kyriazakis, "Deep Learning Pose Estimation for Multi-Cattle Lameness Detection," *SSRN Electron. J.*, pp. 1–19, 2022, https://doi.org/10.2139/ssrn.3990637.
[8] W. Kusakunniran, A. Wiratsudakul, U. Chuachan, S. Kanchanapreechakorn, and T. Imaromkul, "Automatic cattle identification based on fusion of texture features extracted from muzzle images," in *Proceedings of the IEEE International Conference on Industrial Technology*, pp. 1484-1489, 2018. https://doi.org/10.1109/ICIT.2018.8352400.
[9] B. Xu *et al.*, "Evaluation of Deep Learning for Automatic Multi-View Face Detection in Cattle," *Agriculture*, vol. 11, no. 11, pp. 1–15, 2021, https://doi.org/10.3390/agriculture11111062.
[10] S. M. Noe, T. T. Zin, P. Tin, and I. Kobayashi, "Automatic detection of mounting behavior in cattle using semantic segmentation and classification," in *LifeTech IEEE 3rd Global Conference on Life Sciences and Technologies*, pp. 227–228, 2021, https://doi.org/10.1109/LifeTech52111.2021.9391980.
[11] S. M. Noe, T. T. Zin, P. Tin, and I. Kobayashi, "Comparing State-of-the-Art Deep Learning Algorithms for the Automated Detection and Tracking of Black Cattle," *Sensors*, vol. 23, no. 1, pp. 1–20, 2023, https://doi.org/10.3390/s23010532.
[12] Y. Yang, M. Komatsu, T. Ohkawa, and K. Oyama, "Real-Time Cattle Interaction Recognition via Triple-stream Network," *Proc. - 21st IEEE Int. Conf. Mach. Learn. Appl. ICMLA*, pp. 61–68, 2022, https://doi.org/10.1109/ICMLA55696.2022.00016.
[13] N. R. Duraiswami, S. Bhalerao, A. Watni, and C. N. Aher, "Cattle Breed Detection and Categorization Using Image Processing and Machine Learning," in *International Conferrence on Advancements in Smart, Secure, and Intelligent Computing*, pp. 1-6, 2022, https://doi.org/10.1109/ASSIC55218.2022.10088381.
[14] F. Sun, H. Wang, and J. Zhang, "A Recognition Method of Cattle and Sheep Based on Convolutional Neural Network," in *Proceedings 2nd International Seminar on Artificial Intelligence, Networking and Information Technology, AINIT*, pp. 420–424, 2021, https://doi.org/10.1109/AINIT54228.2021.00088.
[15] F. Rozy, F. E. Gani, and A. Dharma, "Comparative Analysis of Convolutional Neural Network Methods in Detecting Mask Wear," *Budapest Int. Res. Critics Institute-Journal*, vol. 5, pp. 16792–16801, 2022, https://doi.org/10.33258/birci.v5i2.5605.
[16] S. Jing, H. Kun, Y. Xin, and H. Juanli, "Optimization of Deep-Learning Network Using Resnet50 Based Model for Corona Virus Disease (COVID-19) Histopathological Image Classification," in *IEEE International Conference on Electrical Engineering, Big Data and Algorithms, EEBDA*, pp. 992–997, 2022, https://doi.org/10.1109/EEBDA53927.2022.9744883.
[17] D. Dagar, Y. Dagar, and R. Sharma, "Face-Mask Recognition and Detection Using Deep Learning," in *International Conference on Machine Learning, Big Data, Cloud and Parallel Computing, COM-IT-CON*, pp. 692–697, 2022, https://doi.org/10.1109/COM-IT-CON54601.2022.9850742.
[18] S. MacHiraju, S. Urolagin, R. K. Mishra, and V. Sharma, "Face Mask Detection using Keras, Opencv and Tensorflow by Implementing Mobilenetv2," in *3rd International Conference on Advances in Computing, Communication Control and Networking, ICAC3N*, pp. 1485–1489, 2021, https://doi.org/10.1109/ICAC3N53548.2021.9725546.
[19] S. A. Sanjaya and S. A. Rakhmawan, "Face Mask Detection Using MobileNetV2 in the Era of COVID-19 Pandemic," in *International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy, ICDABI*, pp. 1–5, 2020, https://doi.org/10.1109/ICDABI51230.2020.9325631.
[20] G. Lakshmi Durga, H. Potluri, A. Vinnakota, N. P. Prativada, and K. C. Yelavarti, "Face Mask Detection using MobileNetV2," in *2nd International Conference on Artificial Intelligence and Smart Energy, ICAIS*, pp. 933–940, 2022, https://doi.org/10.1109/ICAIS53314.2022.9742933.
[21] N. A. Nainan, H. Jeevika, R. Jalan, N. Ramya Shree, C. V. Karthik, and S. P. Shankar, "Real Time Face Mask Detection Using MobileNetV2 and InceptionV3 Models," in *IEEE Mysore Sub Section International Conference, MysuruCon 2021*, pp. 341–345, 2021, https://doi.org/10.1109/MysuruCon52639.2021.9641675.
[22] F. Jia, "Fish Detection Based on Halcon and MobileNetV2 for Automatic Fish Feeding System," in *5th World Conference on Mechanical Engineering and Intelligent Manufacturing, WCMEIM*, pp. 821–824, 2022, , https://doi.org/10.1109/WCMEIM56910.2022.10021553.
[23] S. Kumar, A. Narang, M. Parihar, and V. Sawant, "Breast Tumour Classification using MobileNetV2," in *2021 2nd Global Conference for Advancement in Technology, GCAT*, pp. 1–6, 2021, https://doi.org/10.1109/GCAT52182.2021.9587501.

[24] Y. J. Cheng, W. Lin, Y. Z. Liu, and L. Sun, "Classification of skin diseases based on improved MobileNetV2," in *Proceedings of the 33rd Chinese Control and Decision Conference, CCDC*, pp. 598–603, 2021, https://doi.org/10.1109/CCDC52312.2021.9602387.

[25] T. Adar, E. K. Delice, and O. Delice, "Detection of COVID-19 From A New Dataset Using MobileNetV2 and ResNet101V2 Architectures," *TIPTEKNO Med. Technol. Congr. Proc.*, pp. 1–4, 2022, https://doi.org/10.1109/TIPTEKNO56568.2022.9960225.

[26] K. Dong, C. Zhou, Y. Ruan, and Y. Li, "MobileNetV2 Model for Image Classification," in *2nd International Conference on Information Technology and Computer Application, ITCA*, pp. 476–480, 2020, https://doi.org/10.1109/ITCA52113.2020.00106.

[27] C. R. Yoon and D. H. Kim, "Mobile Convolutional Neural Networks for Facial Expression Recognition," in *International Conference on ICT Convergence*, pp. 1315–1317, 2020, https://doi.org/10.1109/ICTC49870.2020.9289486.

[28] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 4510–4520, 2018, https://doi.org/10.1109/CVPR.2018.00474.

[29] M. B. Hossain, S. M. H. S. Iqbal, M. M. Islam, M. N. Akhtar, and I. H. Sarker, "Transfer learning with fine-tuned deep CNN ResNet50 model for classifying COVID-19 from chest X-ray images," *Informatics Med. Unlocked*, vol. 30, p, 100916, 2022, https://doi.org/10.1016/j.imu.2022.100916.

[30] I. Z. Mukti and D. Biswas, "Transfer Learning Based Plant Diseases Detection Using ResNet50," in *2019 4th International Conference on Electrical Information and Communication Technology, EICT*, pp. 20–22, 2019, https://doi.org/10.1109/EICT48899.2019.9068805.

[31] S. Petluru and P. Singh, "Transfer Learning-based Facial Expression Recognition with modified ResNet50," in *IEEE World Conference on Applied Intelligence and Computing, AIC*, pp. 237–241, 2022, https://doi.org/10.1109/AIC55036.2022.9848956.

[32] Q. Ji, J. Huang, W. He, and Y. Sun, "Optimized deep convolutional neural networks for identification of macular diseases from optical coherence tomography images," *Algorithms*, vol. 12, no. 3, pp. 1–12, 2019, https://doi.org/10.3390/a12030051.

[33] Z. Zhang and Z. Ji, "Open-set Pets Facial Recognition Using Deep Learning and Statistical Learning," in *Proceedings 4th International Conference on Data Intelligence and Security, ICDIS*, pp. 283–287, 2022, https://doi.org/10.1109/ICDIS55630.2022.00050.

[34] S. Xu, Q. He, S. Tao, H. Chen, Y. Chai, and W. Zheng, "Pig Face Recognition Based on Trapezoid Normalized Pixel Difference Feature and Trimmed Mean Attention Mechanism," *IEEE Trans. Instrum. Meas.*, vol. 72, 2023, https://doi.org/10.1109/TIM.2022.3232093.

[35] N. M. Arago *et al.*, "Automated Estrus Detection for Dairy Cattle through Neural Networks and Bounding Box Corner Analysis," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 9, pp. 303–311, 2020, https://doi.org/10.14569/IJACSA.2020.0110935.

[36] Y. Qiao, C. Clark, S. Lomax, H. Kong, D. Su, and S. Sukkarieh, "Automated Individual Cattle Identification Using Video Data: A Unified Deep Learning Architecture Approach," *Front. Anim. Sci.*, vol. 2, pp. 1–14, 2021, https://doi.org/10.3389/fanim.2021.759147.

[37] B. Xu *et al.*, "Automated cattle counting using Mask R-CNN in quadcopter vision system," *Comput. Electron. Agric.*, vol. 171, p. 105300, 2020, https://doi.org/10.1016/j.compag.2020.105300.

[38] A. Yilmaz, G. N. Uzun, M. Z. Gurbuz, and O. Kivrak, "Detection and breed classification of cattle using YOLO v4 algorithm," in *International Conference on INnovations in Intelligent SysTems and Applications, INISTA Proceedings*, pp. 2021–2024, 2021, https://doi.org/10.1109/INISTA52262.2021.9548440.

[39] P. Zhao *et al.*, "A Comparative Study of Deep Learning Classification Methods on a Small Environmental Microorganism Image Dataset (EMDS-6): From Convolutional Neural Networks to Visual Transformers," *Front. Microbiol.*, vol. 13, 2022, https://doi.org/10.3389/fmicb.2022.792166.

[40] A. K. Nugroho, D. M. K. Nugraheni, T. A. Putranto, I. K. E. Purnama, and M. H. Purnomo, "Classification of Ischemic Stroke with Convolutional Neural Network (CNN) approach on b-1000 Diffusion-Weighted (DW) MRI," *Emit. Int. J. Eng. Technol.*, vol. 10, no. 1, pp. 195–216, 2022, https://doi.org/10.24003/emitter.v10i1.694.

[41] S. E. Abdallah, W. M. Elmessery, M. Y. Shams, N. S. A. Al-Sattary, A. A. Abohany, and M. Thabet, "Deep Learning Model Based on ResNet-50 for Beef Quality Classification," *Inf. Sci. Lett.*, vol. 12, no. 1, pp. 289–297, 2023, https://doi.org/10.18576/isl/120124.

[42] A. K. Nugroho, T. A. Putranto, I. K. E. Purnama, and M. H. Purnomo, "Multi Segmentation Method for Hemorraghic Detection," in *International Conference on Intelligent Autonomous Systems, ICoIAS*, pp. 62–66, 2018, https://doi.org/10.1109/ICoIAS.2018.8494039.

[43] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Comput. Electron. Agric.*, vol. 147, pp. 70–90, 2018, https://doi.org/10.1016/j.compag.2018.02.016.