ILKOM Jurnal Ilmiah Vol. 17, No. 2, August 2025, pp. 162-169
Accredited 2nd by RISTEKBRIN No. 10/C/C3/DT.05.00/2025; E-ISSN 2548-7779 | P-ISSN 2087-1716

Deep Learning Convolutional Neural Networks on Multi Label Image Classification of Torajanese Buffalo

Aslan Poetra Ramadhan [a,1,*]; Anik Nur Handayani[a,2]; Ilham Ari Elbaith Zaeni[a,3]
[a] *Department of Electrical Engineering and Informatics, Faculty of Engineering, Universitas Negeri Malang, Jl. Semarang No.5, Kota Malang, 65145, Indonesia*
[1] *aslan.poetra.2105348@students.um.ac.id;* [2] *aniknur.ft@um.ac.id;* [3] *ilham.ari.ft@um.ac.id;* [4] *abdulrachman.manga@umi.ac.id*
* *Corresponding author*

Article history: Received April 28, 2025; Revised July 02, 2025; Accepted August 14, 2025; Available online August 19, 2025

Abstract

Convolutional Neural Networks (CNNs) represent the primary methodology in the advancement of intelligent systems and technologies. The capacity to transition from prediction to categorization establishes CNNs as the primary benchmark in the advancement of deep artificial intelligence. This study use CNN implementation to categorize photos of Torajanese buffalo. The Torajanese buffalo is a distinctive animal species belonging to the Bos bubalis family, integral to the lives and culture of the Torajanese people residing in northern South Sulawesi. This species is integral to the culture, deeply intertwined with several traditional practices of the community. This renders the species distinctive for more investigation. The distinctiveness of the buffalo's style, coloration, and form differentiates them from one another. This study use Convolutional Neural Networks (CNNs) as the primary method to categorize Torajanese buffalo species using head photos and markers derived from local knowledge. This research employs InceptionV3, DenseNet, and Xception as primary architectures, each with variations corresponding to 10, 50, and 100 epochs, therefore enhancing the study. The findings of this investigation indicate that the InceptionV3 architecture has commendable performance across both versions, achieving an average AUC-ROC score of 0.96, although with increased execution time. Nonetheless, the DenseNet architecture demonstrates superior performance in its optimal configuration, achieving flawless accuracy; nonetheless, it incurs the most processing time for the frontal view of the Torajanese buffalo head test case.

Keywords: CNN; Multi Label; Image Processing; Classification; Torajanese Buffalo.

Introduction

The Toraja buffalo, known locally as ‘Tedong’, has unique characteristics that set it apart from other buffalo breeds. In the cultural aspect, buffaloes are very important animals in every traditional ceremony of the Tana Toraja people [1]. The Rambu Solo ‘ceremony (death feast) and Rambu Tuka’ ceremony (thanksgiving feast) have buffaloes as animals that must be offered to the ancestors. The central role of buffaloes in various traditional ceremonies in the life of the Toraja people is also influenced by the social strata of the Toraja people [1], [2]. Torajans believe that buffaloes are the most highly stratified animal among other livestock. Therefore, it has a high social value that affects the economic value of the animal. Economically, the Toraja buffalo is a valuable asset [3], [4].

Toraja buffaloes have several varieties that are distinguished by skin colour and body pattern [5], [6]. The Toraja people divide these varieties of Toraja buffalo into several types. Each buffalo variety has a varied economic level. Generally, the most commonly found Toraja buffaloes are black buffaloes with varied horn shapes [7]. However, in the higher castes of society, especially in the sapu randanan ceremony which is the highest level of rambu solo' ceremony, buffaloes with albino varieties are commonly found in the ceremonies of Torajan indigenous people in this caste [3], [4], [7]. In addition to the albino physical characteristics, the colour of the horns of this type of buffalo with fantastic economic value has a unique colour that is different from the type of buffalo commonly found in traditional ceremonies of the Toraja people [8], [9].

In recent years, the field of image classification has seen rapid progress through the proliferation of deep learning techniques. Among these techniques, the Convolutional Neural Network (CNN) has proven to be very effective for analysing visual data due to its ability to extract important features from images in an efficient manner [10], [11],[12]. Various approaches are being taken to obtain more in-depth information from this digital data [13], [14]. Early recognition of diseases, identification of movements, and identification of plants and animals through image data have been developed along with the development of existing science and technology [15]. Various approaches through a series of researches carried out by previous studies, form special modelling on certain image data [16]. A series of mathematical approaches such as the Convolutional Neural Network (CNN) [17], Support Vector Machine (SVM) [18], [19], Hypergraph Learning (HL) [20], to Extreme Learning Machine (ELM) [21] are a series of approaches used

http://dx.doi.org/10.33096/ilkom.v17i2.2905.162-169

in research related to digital image processing. This approach generally aims to develop a machine that can study the patterns and characteristics of image data to obtain in-depth information from the data [22].

However, studies related to the identification of image objects based on the acquisition of knowledge philosophically through the acquisition of knowledge from local communities have not been found. So many aspects can still be explored further. For example, in the case study of this research, which aims to identify the type of buffalo in the culture of the Tana Toraja people through philosophical knowledge that has a different perspective from science. This research is expected to explore these aspects based on academic publications with the aim of building comprehensive knowledge as a contribution to science in the processing of digital image data for the classification of compound multilabel objects. Identifying and classifying Toraja buffalo appropriately using modern approaches through technology can make a significant contribution to the conservation and management of these genetic resources. In-depth research on the morphological and genetic characteristics of Toraja buffalo not only helps in the conservation of this species but also in increasing comprehensive understanding to maintain the cultural and economic sustainability of the local Toraja community.

Method

This research encompasses multiple phases: data acquisition, data preprocessing, data modeling, data classification, and models performance evaluations The subsequent figure illustrates each of these stages by **Figure 1**.

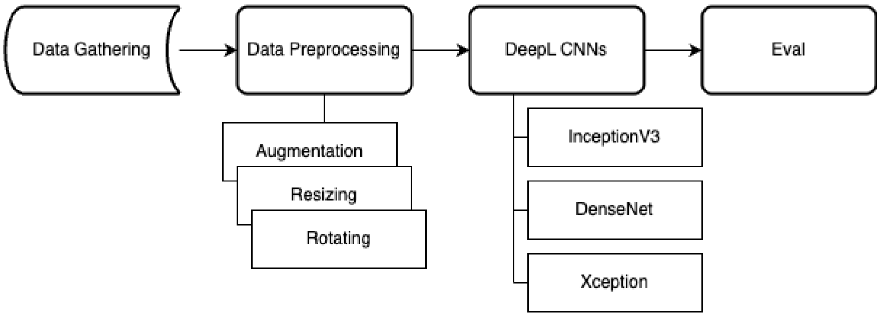


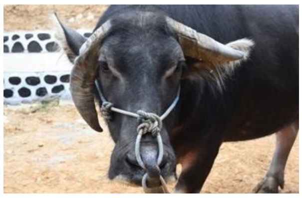
Figure 1. Research Design

A. Data Acquisition

Data Acquisition is a phase carried out with the aim of digitizing the actual conditions of the research subject, allowing the generated digital data to be processed and analyzed using automated methods. These images can be obtained through direct field surveys. The dataset collection involves the use of high-resolution cameras to take frontal photos of Toraja buffalo, resulting in images that clearly display the buffalo's head and horns. The Toraja buffalo are photographed with high-resolution cameras from an average distance of 1.5 meters in clear weather from morning to afternoon to ensure optimal lighting during data collection [23]. Each obtained image is then categorized according to the distinctive characteristics of each Toraja buffalo variety. The labeling method will significantly impact the effectiveness of the CNN model in recognizing and categorizing images according to the assigned labels. A sufficiently large volume of data is also very important at this stage. Increasing the volume of data enhances the model's ability to identify existing patterns and attributes. Therefore, an effective data collection approach will be important in ensuring the availability of adequate data for model training. The dataset is shown in **Figure 2**.


(a) Non-Albino Flat Horn


(b) Non-Albino Curve Horn


(c) Non-Albino Sigmoid Horn


(d) Albino Inverse Horn


(e) Albino Flat Horn


(f) Albino Curve Horn

Figure 2. Toraja Buffalo Classified Label

B. Data Preprocessing

In order to ensure that the visual data utilized for model training is in optimal shape. This procedure encompasses multiple essential stages, including noise elimination, contrast augmentation, high-level compression, and image scaling. Noise removal seeks to eradicate undesirable disturbances in the image, including spots or other artifacts that may hinder analysis [24]. Contrast enhancement is executed to emphasize significant features of the image [25]. Subsequently, high-level compression is applied to the images, as data obtained from high-resolution cameras generates substantial file sizes [21]. This compression employs WebP technology, yielding considerable reductions in file size without sacrificing image resolution. Additionally, resizing the images guarantees uniform dimensions in accordance with the input specifications of the CNN model. This study involves cropping image data from a 16:9 aspect ratio to a 1:1 aspect ratio. Data augmentation constitutes a component of the pre-processing phase. This approach entails altering the source photos to provide supplementary variations, including rotation, flipping, and zooming [26], [27]. The primary dataset is augmented by randomly rotating the photos at varying angles. Furthermore, a flipping procedure is applied to the dataset, generating new data by horizontally inverting the picture data, while data augmentation is executed through zooming, which involves randomly increasing the existing data to create additional data. This efficiently enhances the volume of training data and fortifies the model against variances in data, including discrepancies in perspective or lighting conditions. The outcome of this phase is a compilation of datasets comprising source data and enriched data [28]. The outcomes of this data preprocessing are displayed in **Figure 3**.

(a) Data Source (b) Preprocessing Result

Figure 3. Data Preprocessing Outcome

C. Deep Learning Convolutional Neural Networks

Convolutional Neural Networks utilize the fundamental principles of Neural Network algorithms, incorporating additional layers. CNN employs convolutional layers to process spatial information in images, whereas fully connected layers has the capacity to retain information from multilabel image data [29], [30]. Selecting the appropriate CNN architecture is essential for attaining best outcomes [31]. This study employs the architectures InceptionV3 [32], DenseNet [33], and Xception [34]. The transfer learning method is utilized to reduce training duration and enhance model accuracy, as the pre-trained model has acquired the ability to identify many common features in images [35]. The successful execution of this phase will allow the system to autonomously detect and categorize Toraja buffalo accurately. The architecture of Convolutional Neural Networks (CNNs) is illustrated in **Figure 4**.

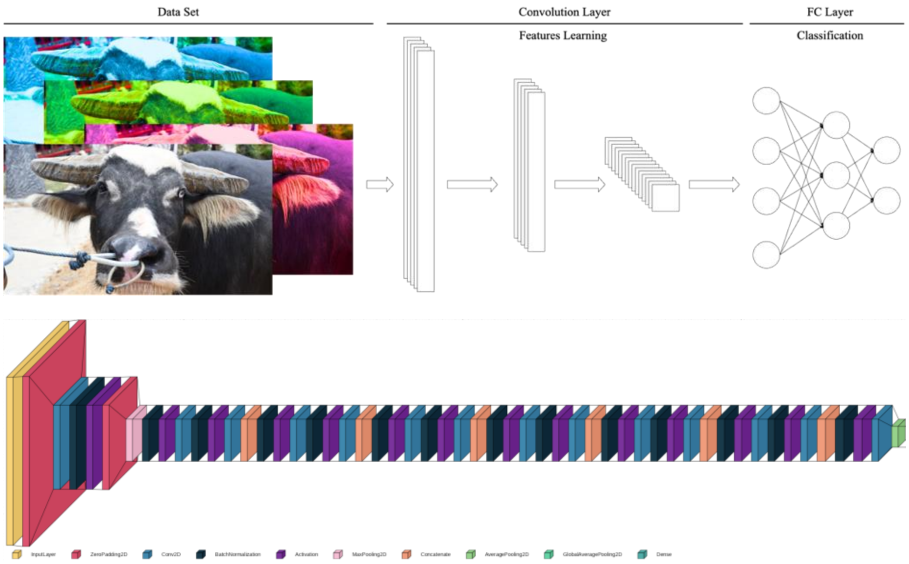


Figure 4. CNNs Architecture & Visualization

D. *Evaluation*

To evaluate the model's efficacy in multilabel image classification, the final stage is model evaluation. It is crucial to conduct this evaluation in order to ensure that the model that has been developed delivers reliable and optimal performance. This study employs accuracy, precision, recall and the AUC-ROC (Area Under the Curve-Receiver Operating Characteristic) evaluation metric to evaluate the model's ability to differentiate between classes. A test dataset that is distinct from the training dataset is used to evaluate the model to prevent it from overfitting [36], [37], [38].

The proportion of all predictions both positive and negative that model got correct defined as accuracy with this formulations below.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

Where TP stands for true positive, TN for true negatives, FP stands for false positive and FN for false negative. Other metrics that we used to measure out of all predictions of a positive class, the percentage that were actually correct are precision that formulated below.

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

Subsequently, evaluation metrics to calculate out of all actual positive sample, the percentage the model correctly identifies or sensitivity of the model are recall with formulation.

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

For plotting the true positive rate against the false positive rate for different detection thresholds and summarize it as a single value that representing probability of the models that ranks a random positve instance higher than a random negative instance with AUC-ROC approach [39].

$$AUC - ROC = \int_0^1 TPR(FPR^{-1}(x))dx \tag{4}$$

Which TPR is True Positive Rate and FPR is False Positive Rate that can be defined with this formulations below.

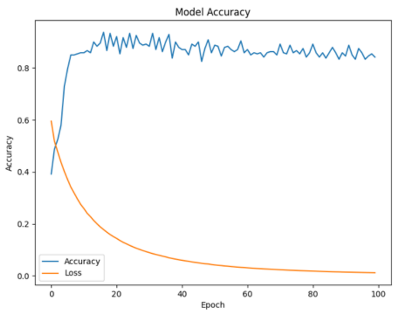
$$TPR = \frac{TP}{P} \tag{5}$$

$$FPR = \frac{FP}{N} \tag{6}$$

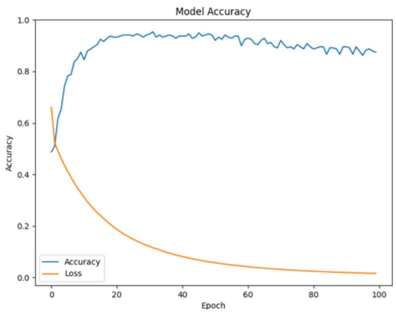
That TP stands for True Positive (number of correct positive predictions), P for Positive, FP for False Positive (number of incorrect positive predicition) and N for Negative.

Results and Discussion

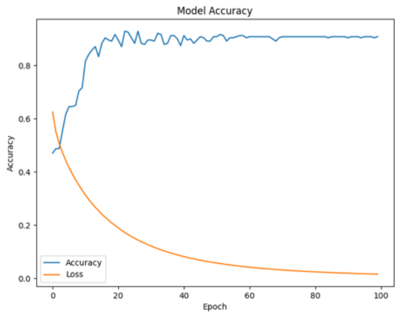
This study yields pre-trained Convolutional Neural Network (CNN) models using several CNN architectures, such as InceptionV3, Xception, and DenseNet. The CNN models were executed under many situations, specifically test scenarios using train-test data ratios of 50:50, 60:40, 70:30, 80:20, and 90:10, alongside model training scenarios utilizing 10 epochs, 50 epochs, and 100 epochs to identify the optimal model from diverse architectures. Multiple iterations of these situations are used to evaluate the efficacy and performance of the CNN model in recognizing study subjects. This research performed performance testing on many architectural kinds, with the performance of each design at the peak point shown in **Figure 5**.



(a) InceptionV3 Architecture



(b) DenseNet Architecture



(c) Xception Architecture

Figure 5. CNNs Architectures Metrics Chart

This research case study presents a graph depicting the training results of the InceptionV3 architecture model, along with a compilation of test scenario tables that demonstrate commendable performance, characterised by increasing accuracy and decreasing loss during the training process. The tables below display each testing phase.

Table 1. *InceptionV3 Testing Result*

Versi Model	*Acuracy*	*Percision*	*Recall*	*AUC-ROC*	*Time(s)*
InceptionV3.5s.10e	*0.08*	*0.95*	*0.34*	*0.92*	*52.53*
InceptionV3.10s.50e	*0.91*	*1*	*0.94*	*1*	*11.20*
InceptionV3.15s.100e	*0.95*	*1*	*0.96*	*1*	*10.38*

The evaluation results table indicates that the InceptionV3 model, using 5 steps and 10 epochs, has a poor accuracy rate of 8% in identifying Toraja buffalo photos. Consequently, the model version exhibits a 95% gain in accuracy after a decrease in batch size during an iteration that enhances both the number of steps and epochs in the training process.

Table 2. *DenseNet Testing Result*

Versi Model	*Acuracy*	*Percision*	*Recall*	*AUC-ROC*	*Time(s)*
DenseNet.5s.10e	*0.03*	*1*	*0.24*	*0.86*	*38.68*
DenseNet.10s.50e	*0.86*	*0.98*	*0.92*	*1*	*17.39*
DenseNet.15s.100e	*1*	*1*	*1*	*1*	*18.13*

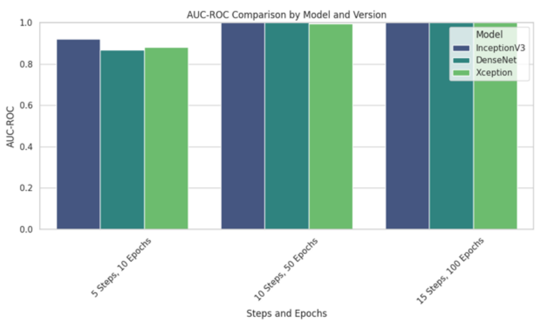
The architecture evaluation table indicates that the DenseNet model version, using 5 steps and 10 epochs, has a poor accuracy of 3% in identifying Toraja buffalo photos. Consequently, the model version exhibits a rise in accuracy to 100% after a decrease in batch size during an iteration that enhances the number of steps and epochs in the training process.

Table 3. *Xception Testing Result*

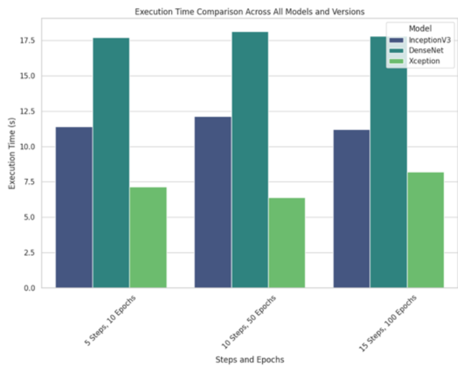
Versi Model	*Acuracy*	*Percision*	*Recall*	*AUC-ROC*	*Time(s)*
Xception.5s.10e	*0.06*	*0.94*	*0.32*	*0.88*	*32.71*
Xception.10s.50e	*0.81*	*1*	*0.88*	*0.99*	*10.54*
Xception.15s.100e	*0.90*	*1*	*0.93*	*0.99*	*7.16*

The evaluation results table indicates that the InceptionV3 model, using 5 steps and 10 epochs, has a poor accuracy rate of 8% in identifying Toraja buffalo photos. Consequently, the model version exhibits a 95% accuracy gain after a decrease in batch size during an iteration that enhances both the number of steps and epochs in the training process.

Figure 6 below compares the performance of the architectures across each version, illustrating that the InceptionV3 architecture in the 5s10e version achieves the highest AUC-ROC value, followed closely by the Xception architecture in the 10s50e version, which exhibits marginally inferior performance compared to the other architectures. In the 15s100e version, all architectures exhibit identical AUC-ROC performance values.



(a) AUC-ROC Comparison by Model and Version



(b) Execution Time Comparison

Figure 6. CNNs Comparison Result

Conclusion

The dataset optimisation procedure is conducted in many phases. The first phase of the dataset experiences a high-level compression procedure that maintains picture resolution using WebP, then converting the images to the JPG format. A cropping procedure with a resolution of 512x512 pixels is executed. The dataset is augmented using several techniques, including rotation, pinching, zooming, and contrast manipulation. The dataset optimisation process occurs in several phases. The first phase of the dataset experiences a high-level compression procedure that maintains picture resolution using WebP, then converting the images to the JPG format. A cropping operation with a resolution of 512x512 pixels is executed. The dataset is augmented using several techniques, including rotation, pinching, zooming, and contrast manipulation. The performance outcomes of multiple evaluated architectures indicate that the InceptionV3 model consistently attains elevated accuracy levels across different iterations, surpassing other architectures, although the top version of the DenseNet architecture exhibits flawless performance.

References

[1] A. Baan, M. D. Girik Allo, and A. A. Patak, “The cultural attitudes of a funeral ritual discourse in the indigenous Torajan, Indonesia,” *Heliyon*, vol. 8, no. 2, p. e08925, Feb. 2022, doi: 10.1016/j.heliyon.2022.e08925.

[2] R. Tangdialla, Y. Lotong, A. L. Matasik, M. Nyhof, O. Pasoloran, and G. O. Mangiwa, “Rambu Solo’ in the Perspective of Toraja Young Generation,” *Rev. Gest. Soc. E Ambient.*, vol. 18, no. 4, p. e06745, May 2024, doi: 10.24857/rgsa.v18n4-122.

[3] R. Dewi, R. Tandu, R. Pasoloran, R. La’biran, and A. Baan, “Tedong (Buffalo): Symbol of Nobility, Humanity, and Entertaiment in Funeral Ceremony in The Indigenous Torajan, Indonesia,” *Int. J. Relig.*, vol. 5, no. 8, Art. no. 8, May 2024, doi: 10.61707/j1qzmj60.

[4] R. D. Haloho, J. Palayukan, A. Setiadi, E. Rianto, and N. Luthfi, “The Feasibility of Business of Buffalo Used in the Traditional Funeral Ceremony (Rambu solo) in West Sulawesi, Indonesia,” *Adv. Anim. Vet. Sci.*, vol. 12, no. 3, 2024, doi: 10.17582/journal.aavs/2024/12.3.523.531.

[5] P. W. Prihandini *et al.*, “Biodiversity of Indonesian indigenous buffalo: First review of the status, challenges, and development opportunities,” *Vet. World*, pp. 2217–2229, Nov. 2023, doi: 10.14202/vetworld.2023.2217-2229.

[6] M. Rusdin, D. D. Solihin, A. Gunawan, C. Talib, and C. Sumantri, “Genetic Variation of Eight Indonesian Swamp-Buffalo Populations Based on Cytochrome b Gene Marker,” *Trop. Anim. Sci. J.*, vol. 43, no. 1, pp. 1–10, Apr. 2020, doi: 10.5398/tasj.2020.43.1.1.

[7] A. S. Putra, P. Hijjang, and M. Tang, “Ethnic Value in Toraja’s Social Ritual: Cultural Education Perspective,” *J. Edusci*, vol. 4, no. 5, pp. 444–451, Jan. 2024, doi: 10.37899/journallaedusci.v4i5.1010.

[8] S. N. Sirajuddin, I. Rasyid, Baharudin, I. Rosada, and Nurhapsa, “Characteristics of Buffalo Traders at the Bolu Animal Market, Toraja Regency, Indonesia,” *J. Adv. Zool.*, vol. 44, no. 2, pp. 143–147, Sep. 2023, doi: 10.17762/jaz.v44i2.207.

[9] A. K. Sendana, F. Rahman, N. Nasmilah, and H. Sahib, “Interpreting socio-cultural values in Toraja: The symbolism of Garonto’ Sura’ in Kada Tomina,” *Edelweiss Appl. Sci. Technol.*, vol. 9, no. 4, pp. 2835–2844, Apr. 2025, doi: 10.55214/25768484.v9i4.6662.

[10] P. Purwono, A. Ma’arif, W. Rahmaniar, H. I. K. Fathurrahman, A. Z. K. Frisky, and Q. M. U. Haq, “Understanding of Convolutional Neural Network (CNN): A Review,” *Int. J. Robot. Control Syst.*, vol. 2, no. 4, pp. 739–748, Jan. 2023, doi: 10.31763/ijrcs.v2i4.888.

[11] L. Alzubaidi *et al.*, “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions,” *J. Big Data*, vol. 8, no. 1, p. 53, Mar. 2021, doi: 10.1186/s40537-021-00444-8.

[12] Herman *et al.*, “A Comperative Study on Efficacy of CNN VGG-16, DenseNet121, ResNet50V2, And EfficientNetB0 in Toraja Carving Classification,” *Indones. J. Data Sci.*, vol. 6, no. 1, Art. no. 1, Mar. 2025, doi: 10.56705/ijodas.v6i1.220.

[13] F. Aljuaydi, M. Zidan, and A. M. Elshewey, “A Deep Learning CNN-GRU-RNN Model for Sustainable Development Prediction in Al-Kharj City,” *Eng. Technol. Appl. Sci. Res.*, vol. 15, no. 1, pp. 20321–20327, Feb. 2025, doi: 10.48084/etasr.9247.

[14] L. Syafie, H. Azis, and F. T. Admojo, “Optimizing Javanese Numeral Recognition Using YOLOv8 Technology: An Approach for Digital Preservation of Cultural Heritage,” *Indones. J. Data Sci.*, vol. 6, no. 1, Art. no. 1, Mar. 2025, doi: 10.56705/ijodas.v6i1.239.

[15] F. Joiya, “OBJECT DETECTION: YOLO VS FASTER R-CNN,” *Int. Res. J. Mod. Eng. Technol. Sci.*, vol. 4, no. 9, Sep. 2022, doi: 10.56726/IRJMETS30226.

[16] H. Azis, Nasrullah, M. Abdullah, S. Ismail, Purnawansyah, and L. Syafie, “A Comparative Study of YOLO Models for Enhanced Vehicle Detection in Complex Aerial Scenarios,” in *2025 19th International Conference on Ubiquitous Information Management and Communication (IMCOM)*, Bangkok, Thailand: IEEE, Jan. 2025, pp. 1–8. doi: 10.1109/IMCOM64595.2025.10857527.

[17] S. Binta Islam, D. Valles, T. J. Hibbitts, W. A. Ryberg, D. K. Walkup, and M. R. J. Forstner, "Animal Species Recognition with Deep Convolutional Neural Networks from Ecological Camera Trap Images," *Animals*, vol. 13, no. 9, p. 1526, May 2023, doi: 10.3390/ani13091526.

[18] Fernanda Januar Pratama, Wikky Fawwaz Al Maki, and Febryanti Sthevanie, "Big Cats Classification Based on Body Covering," *J. RESTI Rekayasa Sist. Dan Teknol. Inf.*, vol. 5, no. 5, pp. 984–991, Oct. 2021, doi: 10.29207/resti.v5i5.3328.

[19] H. Azis, Nirmala, L. Syafie, Herman, F. Fattah, and T. Hasanuddin, "Unveiling Algorithm Classification Excellence: Exploring Calendula and Coreopsis Flower Datasets with Varied Segmentation Techniques," in *2024 18th International Conference on Ubiquitous Information Management and Communication (IMCOM)*, Kuala Lumpur, Malaysia: IEEE, Jan. 2024, pp. 1–7. doi: 10.1109/IMCOM60618.2024.10418246.

[20] L. Wang and S. Deng, "Hypergraph Convolution Network Classification for Hyperspectral and LiDAR Data," *Sensors*, vol. 25, no. 10, p. 3092, May 2025, doi: 10.3390/s25103092.

[21] N. Bacanin, C. Stoean, M. Zivkovic, D. Jovanovic, M. Antonijevic, and D. Mladenovic, "Multi-Swarm Algorithm for Extreme Learning Machine Optimization," *Sensors*, vol. 22, no. 11, p. 4204, May 2022, doi: 10.3390/s22114204.

[22] [22] S. Nakagawa *et al.*, "Rapid literature mapping on the recent use of machine learning for wildlife imagery," *Peer Community J.*, vol. 3, p. e35, Apr. 2023, doi: 10.24072/pcjournal.261.

[23] A. R. Manga, "Data Buffalo Toraja," vol. 2, May 2025, doi: 10.17632/kbft73pdkw.2.

[24] Y. Quan, Y. Chen, Y. Shao, H. Teng, Y. Xu, and H. Ji, "Image denoising using complex-valued deep CNN," *Pattern Recognit.*, vol. 111, p. 107639, Mar. 2021, doi: 10.1016/j.patcog.2020.107639.

[25] M. Upreti, C. Pandey, A. S. Bist, B. Rawat, and M. Hardini, "Convolutional Neural Networks in Medical Image Understanding," *Aptisi Trans. Technopreneurship ATT*, vol. 3, no. 2, pp. 6–12, Sep. 2021, doi: 10.34306/att.v3i2.188.

[26] P. Oza, P. Sharma, S. Patel, F. Adedoyin, and A. Bruno, "Image Augmentation Techniques for Mammogram Analysis," *J. Imaging*, vol. 8, no. 5, p. 141, May 2022, doi: 10.3390/jimaging8050141.

[27] N. E. Khalifa, M. Loey, and S. Mirjalili, "A comprehensive survey of recent trends in deep learning for digital images augmentation," *Artif. Intell. Rev.*, vol. 55, no. 3, pp. 2351–2377, Mar. 2022, doi: 10.1007/s10462-021-10066-4.

[28] M. M. A. Monshi, J. Poon, V. Chung, and F. M. Monshi, "CovidXrayNet: Optimizing data augmentation and CNN hyperparameters for improved COVID-19 detection from CXR," *Comput. Biol. Med.*, vol. 133, p. 104375, Jun. 2021, doi: 10.1016/j.compbiomed.2021.104375.

[29] Y. Liu, H. Pu, and D.-W. Sun, "Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices," *Trends Food Sci. Technol.*, vol. 113, pp. 193–204, Jul. 2021, doi: 10.1016/j.tifs.2021.04.042.

[30] M. Park, D. Q. Tran, S. Lee, and S. Park, "Multilabel Image Classification with Deep Transfer Learning for Decision Support on Wildfire Response," *Remote Sens.*, vol. 13, no. 19, p. 3985, Oct. 2021, doi: 10.3390/rs13193985.

[31] Harlinda, A. Rendi, H. Azis, D. Indra, L. N. Hayati, and N. Kurniati, "Classification of Cia-Cia Letters Using MobileNetV2 and CNN Methods," in *2025 19th International Conference on Ubiquitous Information Management and Communication (IMCOM)*, Jan. 2025, pp. 1–6. doi: 10.1109/IMCOM64595.2025.10857478.

[32] S. P. G, V. S. K, and D. K, "Enhanced Bird Species Image Recognition and Classification using MobileNet and InceptionV3 Transfer learning Architectures," *ELCVIA Electron. Lett. Comput. Vis. Image Anal.*, vol. 24, no. 1, pp. 118–133, May 2025, doi: 10.5565/rev/elcvia.2020.

[33] L. Bothmann *et al.*, "Automated wildlife image classification: An active learning tool for ecological applications," *Ecol. Inform.*, vol. 77, p. 102231, Nov. 2023, doi: 10.1016/j.ecoinf.2023.102231.

[34] W. Shafik, A. Tufail, C. Liyanage De Silva, and R. A. Awg Haji Mohd Apong, "A novel hybrid inception-xception convolutional neural network for efficient plant disease classification and detection," *Sci. Rep.*, vol. 15, no. 1, p. 3936, Jan. 2025, doi: 10.1038/s41598-024-82857-y.

[35] A. Delplanque, S. Foucher, P. Lejeune, J. Linchant, and J. Théau, "Multispecies detection and identification of African mammals in aerial imagery using convolutional neural networks," *Remote Sens. Ecol. Conserv.*, vol. 8, no. 2, pp. 166–179, Apr. 2022, doi: 10.1002/rse2.234.

[36] M. Sivakumar, S. Parthasarathy, and T. Padmapriya, "Trade-off between training and testing ratio in machine learning for medical image processing," *PeerJ Comput. Sci.*, vol. 10, p. e2245, Sep. 2024, doi: 10.7717/peerj-cs.2245.

[37] S. Surono, M. Rivaldi, and N. Irsalinda, "Classification Using U-Net CN on Multi-Resolution CT Scan Image," in *Frontiers in Artificial Intelligence and Applications*, A. J. Tallón-Ballesteros, Ed., IOS Press, 2024. doi: 10.3233/FAIA241412.

[38] K. Namdar, M. A. Haider, and F. Khalvati, “A Modified AUC for Training Convolutional Neural Networks: Taking Confidence Into Account,” *Front. Artif. Intell.*, vol. 4, p. 582928, Nov. 2021, doi: 10.3389/frai.2021.582928.

[39] E. Richardson, R. Trevizani, J. A. Greenbaum, H. Carter, M. Nielsen, and B. Peters, “The receiver operating characteristic curve accurately assesses imbalanced datasets,” *Patterns*, vol. 5, no. 6, p. 100994, Jun. 2024, doi: 10.1016/j.patter.2024.100994.