JAST (Journal of Animal Science and Technology) TITLE PAGE

Upload this completed form to website with submission

ARTICLE INFORMATION	Fill in information in each box below
Article Type	Research article
Article Title (within 20 words without abbreviations)	Deep learning framework for bovine iris segmentation
Running Title (within 10 words)	Detection of Bovine iris based on deep learning
Author	Heemoon Yoon 1, Mira Park 1, Hayoung Lee2, Jisoon Ahn2, Taehyun Lee2 and Sang-Hee Lee 2
Affiliation	1 School of Information Communication and Technology, University of Tasmania, Hobart, 7005, Australia 2 College of Animal Life Sciences, Kangwon National University, Chuncheon, 24341, South Korea
ORCID (for more information, please visit https://orcid.org)	Heemoon Yoon (https://orcid.org/0000-0002-6655-5485) Mira Park (https://orcid.org/0000-0003-0175-8692) Hayoung Lee (https://orcid.org/0009-0003-9946-4541) Jisoon Ahn (https://orcid.org/0009-0005-8168-1091) Taehyun Lee (https://orcid.org/0009-0008-0650-2948) Sang-Hee Lee (https://orcid.org/0000-0001-8725-4174)
Competing interests	No potential conflict of interest relevant to this article was reported.
Funding sources State funding sources (grants, funding sources, equipment, and supplies). Include name and number of grant if available.	This work was supported by the Technology Development Program S3238047 and RS-2023-00223891), funded by the Ministry of SMEs and Startups (MSS, Korea).
Acknowledgements	Not applicable.
Availability of data and material	Upon reasonable request, the datasets of this study can be available from the corresponding author.
Authors' contributions Please specify the authors' role using this form.	Conceptualization: Park M, Lee SH. Data curation: Jang YS, Choi Y. Formal analysis: Park M, Lee SH. Methodology: Yoon H. Software: Yoon H, Lee SH. Validation: Yoon H, Park M, Ahn J. Investigation: Yoon H. Lee T, Lee H. Writing - original draft: Yoon H, Park M. Writing - review & editing: Lee H, Lee T, Ahn J.
Ethics approval and consent to participate	This article does not require IRB/IACUC approval because there are no human and animal participants.

CORRESPONDING AUTHOR CONTACT INFORMATION

For the corresponding author (responsible for correspondence, proofreading, and reprints)	Fill in information in each box below
First name, middle initial, last name	Sang-Hee Lee
Email address – this is where your proofs will be sent	Sang1799@kangwon.ac.kr
Secondary Email address	Knusang1799@gmail.com
Address	College of Animal Life Sciences, Kangwon National University, Chuncheon, South Korea
Cell phone number	+82-10-5366-4179
Office phone number	+82-33-250-8626
Fax number	

Abstract

Iris segmentation is an initial step for identifying the biometrics of animals when establishing a traceability system for livestock. In this study, we propose a deep learning framework for pixel-wise segmentation of bovine iris with a minimized use of annotation labels utilizing the BovineAAEyes80 public dataset. The proposed image segmentation framework encompasses data collection, data preparation, data augmentation selection, training of 15 deep neural network (DNN) models with varying encoder backbones and segmentation decoder DNNs, and evaluation of the models using multiple metrics and graphical segmentation results. This framework aims to provide comprehensive and in-depth information on each model's training and testing outcomes to optimize bovine iris segmentation performance. In the experiment, U-Net with a VGG16 backbone was identified as the optimal combination of encoder and decoder models for the dataset, achieving an accuracy and dice coefficient score of 99.50% and 98.35%, respectively. Notably, the selected model accurately segmented even corrupted images without proper annotation data. This study contributes to the advancement of iris segmentation and the establishment of a reliable DNN training framework.

Keywords: Cow; Deep learning; Identification; Iris; Segmentation

Introduction

Accurate animal identification applies to individual management and the entire process of livestock food production; hence, it is essential for establishing a traceability system for the food supply chain from farm to table [1, 2]. Reliable animal identification methodologies monitor each stage of growth steps and production while minimizing trade losses and ensuring animal ownership. To implement such a tracking system, a robust identification methodology is required [3] because the failure of the tracking system can cause enormous damage. The damage is linked to cow health and food safety, which can put the health of customers at risk and cause serious economic problems [4].

To eliminate these potential hazards, ear notching, tattoos, tags, and branding are some of the traditional permanent methods used for animal identification. However, these can be easily duplicated, simplifying theft and fraud [5]. Radio frequency identification (RFID) tags have been developed as an alternative to traditional methods [6]. Through RFID, animals are registered in computer systems and can be identified by scanning the RFID tag. However, the tag is invasive and can be changed by manipulating it in the system, creating an avenue for fraud [7]. Recently, biometrics such as retinal vascular patterns (RVPs) [8], muzzle [9, 10], and iris [11, 12] have been proposed to resolve the problems of RFIDs. Methods that utilize these biometrics are reliable for identifying an entity because they are the most accurate and stable biometric modalities during the lifetime of an animal [3, 13].

With the advent of deep neural networks (DNNs) [14], there have been several attempts to identify anatomical parts of an animal using deep learning technologies [10, 15-17]. Among the deep learning technologies, the segmentation technique classifies objects within a given image in a pixel-wise manner. As segmentation of the iris from the image of an eye is essential for initiating iris identification, using an elaborate and accurate segmentation technique is key to successful iris recognition [16].

In this study, we discuss bovine iris segmentation using a novel framework. The framework develops multiple segmentation models by training on publicly available bovine iris datasets, BovineAAEyes80 [18] and comparing combinations of state-of-the-art deep learning techniques. Since iris datasets are rare and have limited formats, like other biometric datasets, we propose a framework that can be used to develop models using the smallest input datasets: region of interests (ROIs) labels and RGB images. This study contributes to the advancement of iris identification using DNNs and the development of a reliable DNN training framework that assists in identifying the most suitable combination of DNN models for biometric images.

Materials and Methods

Framework Overview

The proposed framework starts with data collection. The input data must contain pairs of image and annotation data (Figure 1a). After collecting the data pairs, the data is prepared, which includes data splitting and augmentation selection. Data must be split into training, validation, and test datasets that are preferably mutually exclusive for each of the training, validation, and testing stages to be conducted with unseen data. The data must be split such that it is equally distributed in terms of quality since this step can affect the result of the trained model [19]. The augmentation selection step can be varied according to the traits of the dataset (Figure 1b). After selecting the augmentation options, we developed 15 combinations of deep neural network (DNN) models by utilizing three different encoder backbones, namely VGG16 [20], ResNet50 [21], and MobileNet [22]. Additionally, we employed five segmentation decoder DNNs, namely FCN8, FCN16, FCN32 [23], U-Net [24], and SegNet [25]. The encoder and decoder form an architecture known as an encoder-decoder network, which is widely used for tasks such as image segmentation. The encoder extracts useful features and compresses the input data, while the decoder reconstructs or segments the data based on the encoded representation. This architecture enables the network to learn and leverage hierarchical and contextual information, leading to more accurate segmentation results. These combinations allowed us to explore a range of model architectures and evaluate their performance. In total, we trained and evaluated 75 models (15 combinations x 5-fold cross-validation) to ensure the reliability of the training results (Table 1). The evaluation process included assessing various metrics such as accuracy, precision, recall, intersection over union (IoU), and dice coefficient [31]. Furthermore, the framework provided detailed information such as inference time on each model, along with graphical representations of the segmentation results (Figure 1d).

Model Training Environment and Configuration

With reference to previous studies, we compared five candidates, FCN32, FCN16, FCN8, U-Net, and SegNet, to find the most reliable architecture for anatomical segmentation (Figure 1c). All configurations were set to be equal for a fair comparison, minimizing variants between model training processes. After several attempts, the training hyperparameters were experimentally determined: training for 100 epochs with 128 steps per epoch, a learning rate of 0.001 optimized using an Adam optimizer, and a batch size of 4. These trained models automatically generated anatomical ROIs from input test images. After training and evaluation with statistical performance measures, such as the dice coefficient and accuracy [26], the statistical result is returned in the csv format and analyzed within the framework system.

Model training was conducted on Anaconda 4.10.1 running on 64bit Ubuntu Linux 20.04.3 LTS and Python v3.8.8. TensorFlow-GPU v2.7.0 and CUDA 11.4 were used to accelerate the DNNs framework's training process on a 24 GB RTX 3090 graphics card, and Keras v2.7.0 was used as a Python deep learning application programming interface (API). In the BovineAAEyes80 dataset, brightness ±10 and rotation ±

40° augmentation are applied to cover variations that could arise in from the capturing environment, such as non-cooperative behavior of bovines and changes in lighting conditions [18].

Model Evaluation

The classification performance of the trained model was evaluated using the following metrics: accuracy (1), recall (2), precision (3), IoU (4), and dice coefficient (5) [26]. Compared to the reference annotation, each pixel is classified into one of four outcomes: true positive (TP), true negative (TN), false positive (FP), and false negative (FN); these classifications are according to the metric criteria of a previous study [27].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1) \qquad Recall = \frac{TP}{TP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3) \qquad IoU = \frac{TP}{TP + FP + FN} \quad (4)$$

$$Dice = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (5)$$

Results and discussion

The learning curves of model training, as shown in Figure 2, provide important insights into the performance and stability of different models during the training process. In the training curve of VGG16, as seen in Figures 2a and 2b, all of the FCN series are observed to be unstable during the training process. Additionally, FCN32 is found to have the highest loss and the lowest accuracy, indicating that it is not the best model for this particular task. On the other hand, SegNet and U-Net demonstrate a comparatively stable decrease in loss and increase in accuracy during most of the training process. In the training curves of ResNet50, as depicted in Figures 2c and 2d, and MobileNet, as seen in Figures 2e and 2f, decent accuracies and losses with little fluctuation, compared with VGG16, are observed. In contrast to FCN32, which has the poorest performance among the models, the other models show promising results.

Table 2 shows the test results of the models trained with an unseen test dataset. In Table 2, U-Net with a MobileNet backbone has the best dice coefficient (98.35 ± 0.54%), accuracy (99.50 ± 0.16%), and precision (99.57 ± 0.16%). U-Net with a VGG16 backbone shows the best IoU score (96.81 ± 2.01%), which is slightly (0.01%) better than that of U-Net with a MobileNet backbone.

Table 3 presents the inference times of different decoder and encoder models for a given task. While MobileNet is generally observed to perform the fastest across most of the decoder models, the performance of different decoder and encoder model combinations can be influenced by a range of factors beyond the choice of encoder architecture alone. For instance, when paired with FCN8 and FCN16, MobileNet has processing times that are slower than those of VGG16 and ResNet50. Specifically, when paired with FCN8, the mean processing times are 133.3 ± 1.0 ms for VGG16, 180.1 ± 1.6 ms for ResNet50, and 156.8 ± 7.1 ms for MobileNet. When paired with FCN16, the mean processing times are 131.6 ± 0.6 ms for VGG16, 182.5 ± 1.3 ms for ResNet50, and 136.7 ± 1.5 ms for MobileNet. Likewise, while MobileNet performs well

when paired with FCN32, it is outperformed by VGG16 when paired with FCN8 and FCN16. However, when MobileNet is paired with SegNet and U-Net, it shows the fastest inference speed recording 122.8 ± 3.2 ms and 116.3 ± 2.5 ms respectively.

These findings suggest that the performance of different decoder and encoder model combinations can be influenced by a range of factors beyond the general performance of encoder architecture alone. The characteristics and complexity of the dataset, as well as the specifics of the task at hand, can all impact the performance of the model. Therefore, it is important to carefully consider the selection of both the decoder and encoder architectures when developing deep learning models for image segmentation tasks. Overall, Table 3 provides useful information on the performance of different decoder and encoder model combinations for the given task, with certain models performing significantly faster or slower than others. The information on processing times can be used to select the optimal model combination based on the trade-off between processing speed not only segmentation accuracy.

Based on the results of our study, the U-Net model with a MobileNet backbone can be considered the most appropriate model for the given dataset. However, it is important to note that there are significant variations in the size of a segmentation unit of pixels between different backbones, which is influenced by the extracted feature map size of each encoder architecture. Therefore, when selecting a model, both numerical scores and pixel segmentation size should be taken into account, as the optimal DNN model can vary depending on the application domain.

In the context of iris segmentation, where the fine segmentation of the edge of the iris boundaries is a targeted objective, the model with the second-best score, U-Net with a VGG16 backbone, was chosen as the best model due to its superior dense boundary segmentation. The decision to select this model was based on the median values of the dice coefficient observed in the results of the 5-fold cross-validation.

Overall, while the U-Net model with a MobileNet backbone is the most suitable for the given dataset, the U-Net model with a VGG16 backbone was deemed the optimal model for iris segmentation due to its superior boundary segmentation. The selection of the best model for a given task requires careful consideration of both numerical scores and pixel segmentation size, as well as the specific objectives of the application domain.

In Figure 4, common corruptions in iris images are described. Minor corruptions, which distort the iris image, can be caused by many factors, such as dust in the spots, stains over the lens, the animal's eyelash or fur in the eye (Figure 4a), and unwanted light spots [16]. These minor corruptions were not reflected in the segmentation result (Figure 4c). However, this issue must be resolved to eliminate false information within the iris image. Major corruption is generally caused by relatively large parts of the animal's body, such as the occlusion of eyelashes and eyelids (Figure 4b). As mentioned in other studies, these major corruptions can impede accurate identification [12, 18, 28]. However, the best selected model accurately segmented the corrupted image by excluding the occlusion (Figure 4d). This could have not been calculated

correctly in the result because the annotation labels used in the model training did not provide much pixel-wise accurate segmentation ground truth. This is remarkable compared with other studies using image processing techniques because it segmented the exact iris area without preprocessing or postprocessing with the model's knowledge, even though image corruption was not annotated in the given labels.

The field of deep learning is rapidly evolving, with new and improved models being developed all the time [29, 30, 31]. Therefore, it is possible that even better-performing segmentation encoder and feature map decoder models may become available in the future. The current study used a limited set of models, which may not represent the best possible models for bovine iris segmentation. However, the proposed deep learning framework provides a foundation for future research to incorporate and evaluate additional models. This could lead to further improvements in the accuracy and efficiency of the segmentation process.

In addition, the present study focused on bovine iris segmentation using a limited dataset. Future research could expand the framework to include other animal species and biometric features. This would increase the framework's versatility and applicability to various animal biometric applications.

Overall, while the proposed framework has limitations, it serves as a starting point for future research to incorporate additional models and further optimize the segmentation process. As the field of deep learning continues to advance, it holds great promise for improving animal identification and traceability systems.

The deep learning framework proposed in this study for bovine iris segmentation has potential applications in animal identification and traceability systems, which are crucial for ensuring food safety, quality, and individual management system. The framework could be used in various animal biometric applications, such as identifying individual animals in large herds, monitoring animal health, and tracking animal movements.

In addition, the proposed framework could have implications for improving the efficiency and accuracy of livestock management practices. By enabling reliable and rapid animal identification, the framework could help reduce labor costs and improve animal welfare. Furthermore, the framework's use of deep learning technology could lead to new insights into animal biometrics and behavior, which could inform the development of more effective management strategies. Moreover, the proposed framework's reliance on deep learning technology could also exacerbate existing biases and inequalities in animal identification and traceability systems. Careful consideration must be given to how the framework's use of biometric data might disproportionately affect certain animal populations or communities. Nevertheless, despite these concerns, it is deemed necessary to conduct in-depth further research as this technology can still contribute to the national animal population management system, livestock distribution industry, and livestock quality assessment.

In summary, the proposed deep learning framework for bovine iris segmentation has the potential to improve animal identification and traceability systems, and to enhance the efficiency and accuracy of

livestock management practices. However, its use must be guided by ethical principles and considerations to prevent potential harms and biases.

Conclusion

With the proposed framework, iris segmentation for identifying animal biometrics was performed utilizing the information in the trained DNNs along with robust comparisons to determine the best model for the given dataset. The model selected as the best combination of an encoder and decoder, U-Net with a VGG16 backbone, demonstrated an accuracy and dice coefficient of 99.50% and 98.35%, respectively, on an unseen test dataset.

This study contributes to the initial step of iris identification to improve animal tracking systems; it suggests a framework for training DNNs for pixel-wise segmentation using a minimum use of annotation labels. For the reliable comparison of various combinations of DNN models to select the most suitable DNN model combination, this approach uses multiple metrics commonly used in the evaluation of segmentation, including visual references; hence, it is unbiased and has consistent model selection. The framework has the potential to improve the accessibility of DNNs for operators with limited knowledge of DNNs, accelerate inter-study comparisons, and reduce the variations in current manual model selection methods. Following this study, the authors plan to improve the framework's model selection, image segmentation, machine learning, animal biometrics, and multi-resolution imaging. The goal of future research is to develop techniques and skills that can be applied to animal tracking, image recognition, and artificial intelligence applications in domestic animal field.

References

1. Eradus WJ, Jansen MB. Animal identification and monitoring. Comp Electron Agric. 1999;24:91-98.

2. Pendell DL, Brester GW, Schroeder TC, Dhuyvetter KC, Tonsor GT. Animal identification and tracking in the United States. Am J Agric Econ. 2010;92:927-940. https://doi.org/10.1093/ajae/aaq037

3. Awad AI. From classical methods to animal biometrics: A review on cattle identification and tracking. Comp Electron Agric 2016;123:423-435. https://doi.org/10.1016/j.compag.2016.03.01

4. Caporale V, Giovannini A, Francesco CD, Calistri P. Importance of the traceability of animals and animal products in epidemiology. Rev Sci Tech. 2001;20:372-378. https://doi.org/10.20506/rst.20.2.1279

5. Klindtworth M, Wendl G, Klindtworth K, Pirkelmann H. Electronic identification of cattle with injectable transponders. Comp Electron Agric. 1999;24:65-79. https://doi.org/10.1016/S0168-1699(99)00037-X

6. Roberts CM. Radio frequency identification (RFID). Comput Secur. 2006;25:18-26. https://doi.org/10.1016/j.cose.2005.12.003

7. Ruiz-Garcia L, Lunadei L. The role of RFID in agriculture: Applications, limitations and challenges. Comp Electron Agric. 2011;79:42-50. https://doi.org/10.1016/j.compag.2011.08.010

8. Allen A, Golden B, Taylor M, Patterson D, Henriksen D, Skuce R. Evaluation of retinal imaging technology for the biometric identification of bovine animals in Northern Ireland. Livest Sci. 2008;116:42-52. https://doi.org/10.1016/j.livsci.2007.08.018

9. Awad AI, Zawbaa HM, Mahmoud HA, Nabi EHHA, Fayed RH, Hassanien AE. A robust cattle identification scheme using muzzle print images. In: Proceeding of Federated Conference on Computer Science and Information Systems; p. 529-534; 2013; Krakow, Poland.

10. Kumar S, Pandey A, Satwik KSR, Kumar S, Singh SK, Singh AK, Mohan A. Deep learning framework for recognition of cattle using muzzle point image pattern. Measurement 2018;116:1-17. https://doi.org/10.1016/j.measurement.2017.10.064

11. Larregui JI, Espinosa J, Ganuza ML, Castro SM. Biometric Iris Identification in Bovines. In Computer Science & Technology Series; Editorial de la Universidad Nacional de La Plata: Buenos Aires, Argentina, 2015; p. 111-121.

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

13. Daugman J. The importance of being random: statistical principles of iris recognition. Pattern Recognit. 2003;36:279-291. https://doi.org/10.1016/S0031-3203(02)00030-4

14. Miikkulainen R, Liang J, Meyerson E, Rawal A, Fink D, Francon O, Raju B, Shahrzad H, Navruzyan, A, Duffy N, Hodjat B. Evolving deep neural networks. In : Robert Kozma R, Alippi C, Choe Y, Morabito F, editors. Artificial intelligence in the age of neural networks and brain computing, 1st ed.; Amsterdam: Elsevier; 2019. p. 293-312.

15. Al-Waisy AS, Qahwaji R, Ipson S, Al-Fahdawi S, Nagem TA. A multi-biometric iris recognition system based on a deep learning approach. Pattern Anal Appl. 2018;21:783-802. https://doi.org/10.1007/s10044-017-0656-1

16. Arsalan M, Hong HG, Naqvi RA, Lee MB, Kim MC, Kim DS, Kim CS, Park KR. Deep learning-based iris segmentation for iris recognition in visible light environment. Symmetry 2017;9:263. https://doi.org/10.3390/sym9110263

17. Yoon H, Park M, Yeom S, Kirkcaldie MT, Summons P, Lee SH. Automatic Detection of Amyloid Beta Plaques in Somatosensory Cortex of an Alzheimer's Disease Mouse Using Deep Learning. IEEE Access. 2021;9:161926-161936. https://doi.org/10.1109/ACCESS.2021.3132401

18. Larregui JI, Cazzato D, Castro SM. An image processing pipeline to segment iris for unconstrained cow identification system. Open Comput Sci 2019;9:145-159. https://doi.org/10.1515/comp-2019-0010

19. Wu Z, Ramsundar B, Feinberg EN, Gomes J, Geniesse C, Pappu AS, Leswingd K, Pande V. MoleculeNet: a benchmark for molecular machine learning. Chem Sci J. 2018;9:513-530. https://doi.org/10.1039/C7SC02664A

20. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: Proceedings of 3rd International Conference on Learning Representations; 2015; San Diego, CA.

21. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016; Las Vegas, NV.

22. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861 2017. https://doi.org/10.48550/arXiv.1704.04861

23. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2015; Boston, MA.

24. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Proceedings of the International Conference on Medical image computing and computer-assisted intervention; 2015; Munich, Germany.

25. Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal Mach Intell. 2017;39:2481-2495. https://doi.org/10.48550/arXiv.1511.00561

26. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC Med imaging. 2015;15:1-28. https://doi.org/10.1186/s12880-015-0068-x

27. Olson DL, Delen, D. Performance Evaluation for Predictive Modeling. In Advanced data mining techniques; Springer: Berlin, Heidelberg, 2008; pp. 137-147. https://doi.org/10.1007/978-3-540-76917-0_9

28. Cui J, Wang Y, Tan T, Ma L, Sun Z. A fast and robust iris localization method based on texture segmentation. In: Proceedings of the Volume Biometric Technology for Human Identification of the SPIE; Orlando, FL.

29. Wang D, Cao W, Zhang F, Li Z, Xu S, Wu X. A review of deep learning in multiscale agricultural sensing. Remote Sensing. 2022;14:559. https://doi.org/10.3390/rs14030559

30. Garcia R, Aguilar J, Toro M, Pinto A, Rodriguez P. A systematic literature review on the use of machine learning in precision livestock farming. Comput Electro Agri. 2020;179:105826. https://doi.org/10.1016/j.compag.2020.105826

31. Neethirajan S. The role of sensors, big data and machine learning in modern animal farming. Sensing and Bio-Sensing Research, 2020;29:100367.

Tables and Figures

Table 1. Distribution of the dataset for 5-fold cross validation

Dataset	Fold	Images	Eye ID
Train	Fold 1	12	3, 7
	Fold 2	13	4
	Fold 3	13	5, 6
	Fold 4	12	8, 9
	Fold 5	22	10, 11
Test	-	8	1, 2
Total		80 images	11 eyes

Table 2. Test result of the trained models with unseen test dataset

Decoder	Encoder	Dice	IoU	Accuracy	Recall	Precision
FCN8	VGG16	97.90 ± 0.25^{ab}	95.97 ± 0.45^{ab}	99.37 ± 0.07^{a}	96.81 ± 0.44^{ab}	99.06 ± 0.12^{ab}
	ResNet50	98.31 ± 0.44^{a}	96.75 ± 0.81^{a}	99.49 ± 0.13^{a}	97.50 ± 0.76^{a}	99.17 ± 0.09^{a}
	MobileNet	97.14 ± 0.16^{abc}	94.57 ± 0.29^{abc}	99.15 ± 0.04^{a}	95.52 ± 0.37^{abc}	98.91 ± 0.17^{ab}
FCN16	VGG16	96.91 ± 0.41^{abc}	94.17 ± 0.72^{abc}	99.08 ± 0.12^{a}	95.45 ± 0.63^{abc}	98.48 ± 0.28^{abc}
	ResNet50	96.38 ± 1.20^{bc}	93.39 ± 2.02^{bc}	98.96 ± 0.32^{a}	94.46 ± 1.92^{abcd}	98.67 ± 0.22^{ab}
	MobileNet	94.44 ± 0.19^{de}	89.93 ± 0.31^{de}	98.39 ± 0.05^{a}	91.90 ± 0.53^{de}	97.40 ± 0.49^{c}
FCN32	VGG16	89.70 ± 0.21^{f}	81.77 ± 0.25^{f}	93.96 ± 1.07^{c}	88.73 ± 1.20^{f}	91.13 ± 1.30^{e}
	ResNet50	94.00 ± 0.51^{e}	89.23 ± 0.83^{e}	98.29 ± 0.13^{ab}	90.82 ± 0.77^{ef}	97.81 ± 0.36^{bc}
	MobileNet	88.51 ± 0.86^{f}	81.11 ± 1.16^{f}	96.96 ± 0.18^{b}	83.61 ± 1.18^{g}	95.43 ± 0.43^{d}
SegNet	VGG16	96.45 ± 0.75^{bc}	93.40 ± 1.29^{bc}	98.97 ± 0.20^{a}	94.04 ± 1.21^{bcd}	99.24 ± 0.17^{a}
	ResNet50	98.04 ± 0.54^{ab}	96.25 ± 1.01^{ab}	98.05 ± 1.06^{ab}	96.85 ± 1.10^{ab}	99.36 ± 0.13^{a}
	MobileNet	95.73 ± 0.33^{dc}	92.09 ± 0.56^{cd}	98.77 ± 0.09^{a}	92.83 ± 0.61^{cde}	99.12 ± 0.09^{ab}
U-Net	VGG16	98.34 ± 0.49^{a}	96.81 ± 0.90^{a}	99.47 ± 0.81^{a}	97.38 ± 0.83^{a}	99.37 ± 0.13^{a}
	ResNet50	98.26 ± 0.42^{a}	96.66 ± 0.78^{a}	99.18 ± 0.27^{a}	97.11 ± 0.82^{a}	99.52 ± 0.09^{a}
	MobileNet	98.35 ± 0.24^{a}	96.80 ± 0.45^{a}	99.50 ± 0.07^{a}	97.20 ± 0.47^{a}	99.57 ± 0.07^{a}

Dice: Dice coefficient, IoU: Intersection of Union, the values are represented as mean ± standard error mean ($P<0.05$)

Table 3. Inference times of the trained models with unseen test dataset

Decoder	Encoder	Times (ms)		
		Mean	Min.	Max
FCN8	VGG16	133.3 ± 1.0	126.9	151.1
	ResNet50	180.1 ± 1.6	167.8	211.8
	MobileNet	156.8 ± 7.1	124.8	272.1
FCN16	VGG16	131.6 ± 0.6	127.3	144.4
	ResNet50	182.5 ± 1.3	168.2	202.7
	MobileNet	136.7 ± 1.5	125.6	158.7
FCN32	VGG16	135.1 ± 1.5	127.5	165
	ResNet50	184.5 ± 1.5	169.2	207.5
	MobileNet	132.8 ± 2.8	125.6	158.2
SegNet	VGG16	129.2 ± 3.5	107.9	182.3
	ResNet50	126.5 ± 2.1	113.2	168.6
	MobileNet	122.8 ± 3.2	102.8	157.8
U-Net	VGG16	125.6 ± 2.1	108.2	177.8
	ResNet50	143.3 ± 7.8	113.4	347.5
	MobileNet	116.3 ± 2.5	100.8	152.8

The values are represented as mean ± standard error mean.

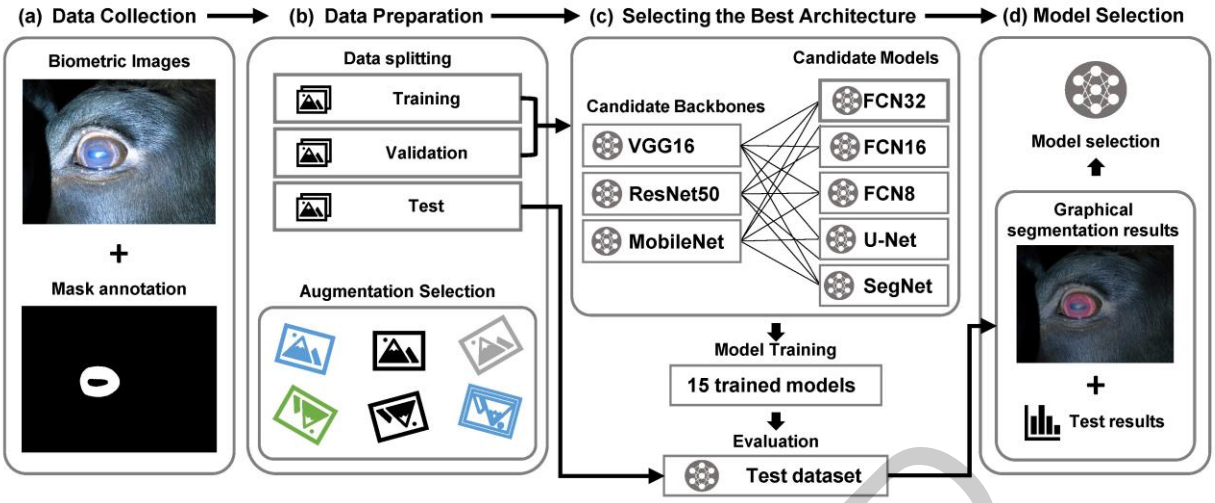

Figure 1. Scheme of model training for selection of the best combination of segmentation models for biometric images. (a) Biometric images are collected and captured. The images' mask anno-tation data is created or collected from data sources. (b) Images are split into training, validation, and test datasets. Augmentation techniques can be selected and adapted within the framework according to the traits of the image. (c) Five DNNs – FCN32, FCN16, FCN8, U-Net and SegNet – are trained and compared with 3 different backbones for each training to select the most reliable model. After training 15 combination models, evaluation is conducted with unseen test dataset.

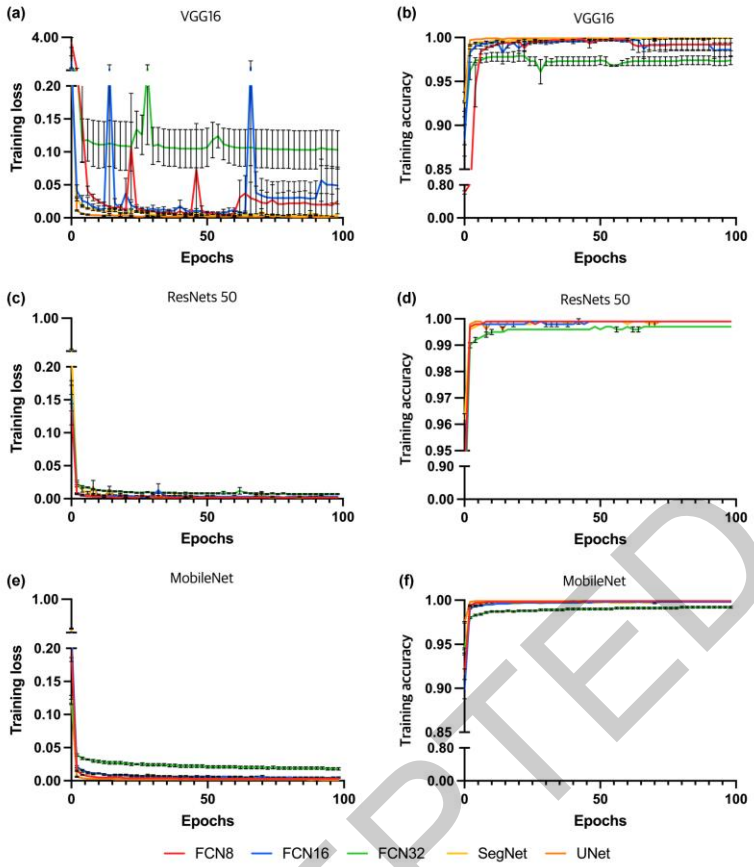

Figure 2. Learning curves of model training

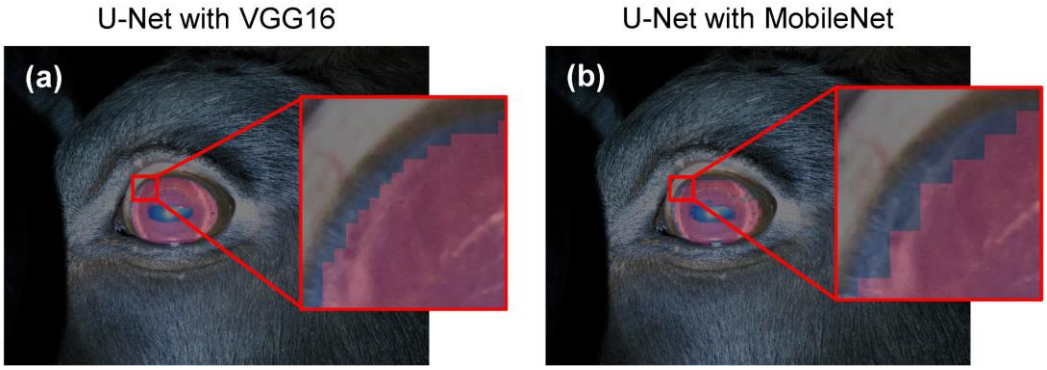

Figure 3. Comparison of segmentation performance in the unit of pixels. Depending on the type of encoder backbone, the size of a segmentation unit is changed because of the DNNs structured and it influences to the model performance. The smaller size of a segmentation unit (a) allows more precise iris boundary segmentation and yields faster inference speed compared to (b).

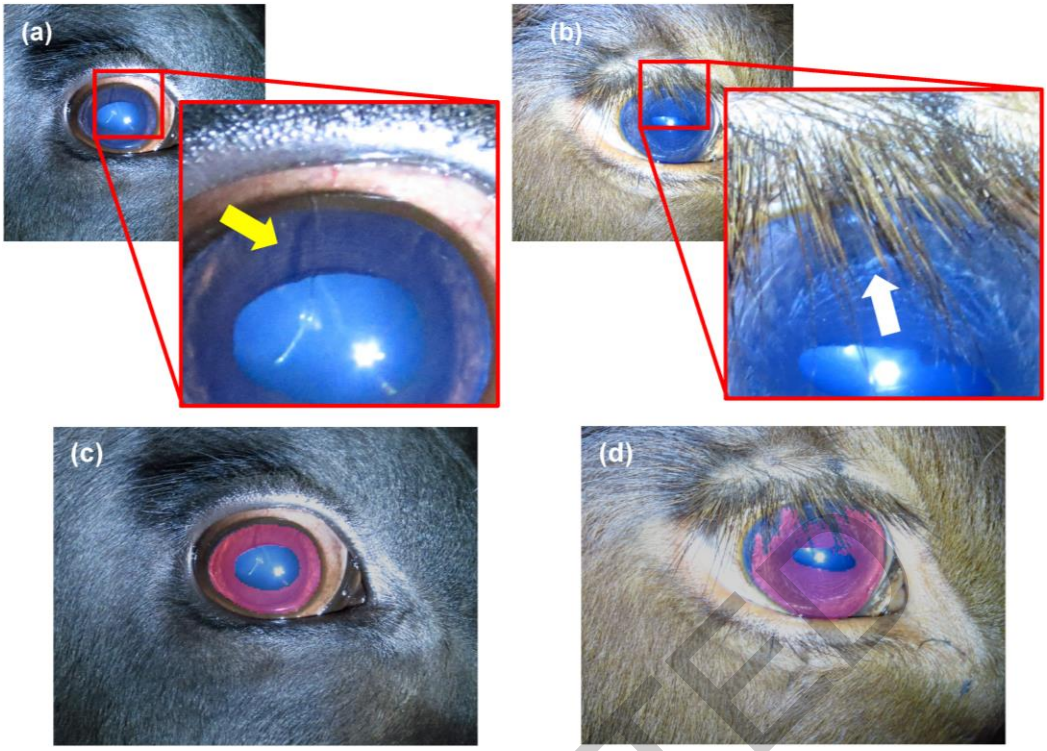

Figure 4. Common corruptions in iris image. (a) Minor corruption caused by an eyelash over the eye. The eyelash made shadow in iris area (yellow arrow). In addition, the eyelash itself covers minor part of iris and pupil. (b) Major corruption caused by eyelashes which cover most part of upper iris. The corruption makes it difficult to identify the area. (c) Segmentation result does not reflect the minor corruption in the iris. (d) Even though the annotation masks used in training the model does not exclude eyelashes from the iris area, the trained model successfully segmented iris area excluding eyelashes.