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-based Gestational Sac Detection in Ultrasound Images using modified YOLOv7-E6E Model
Running Title (within 10 words)	Artificial intelligence in swine husbandry
Author	Tae-kyeong Kim 1, Jin Soo Kim 2, Hyun-chong Cho 3
Affiliation	1 Interdisciplinary Graduate Program for BIT Medical Convergence, Kangwon National University, Chuncheon 24341, Korea 2 College of Animal Life Sciences, Kangwon National University, Chuncheon 24341, Korea. 3 Department of Electronics Engineering and Interdisciplinary Graduate Program for BIT Medical Convergence, Kangwon National University, Chuncheon 24341, Korea
ORCID (for more information, please visit https://orcid.org)	Tae-kyeong Kim (https://orcid.org/0000-0001-8896-416X) Jin Soo Kim (https://orcid.org/0000-0002-9518-7917) Hyun-chong Cho (https://orcid.org/0000-0003-2122-468X)
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 research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2022R1I1A3053872)
Acknowledgements	This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2022R1I1A3053872) and this study was supported by 2022 the RDA Fellowship Program of National Institute of Animal Science, Rural Development Administration, Korea.
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: HCC, JSK Data curation: JSK Formal analysis: TKK, HCC Methodology: TKK, HCC Software: TKK, HCC Validation: JSK, HCC Investigation: TKK, HCC Writing - original draft: TKK, JSK, HCC
Ethics approval and consent to participate	This study was approved by IACUC of Rural Development Administration (No. NIAS-2022-563), Korea.

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	Jin Soo Kim // Hyun-chong Cho
Email address – this is where your proofs will be sent	kjs896@kangwon.ac.kr // hyuncho@kangwon.ac.kr
Secondary Email address	
Address	College of Animal Life Sciences, Kangwon National University, Chuncheon 24341, Korea // Department of Electronics Engineering and Interdisciplinary Graduate Program for BIT Medical Convergence, Kangwon National University, Chuncheon 24341, Korea
Cell phone number	+82-10-2566-5961 // +82-10-5775-8370
Office phone number	+82-33-250-8614 // +82-33-250-6301
Fax number	+82-33-259-5572 // +82-33-259-5674

Abstract

As the population and income levels rise, meat consumption steadily increases annually. However, the number of farms and farmers producing meat decrease during the same period, reducing meat sufficiency. Information and Communications Technology (ICT) has begun to be applied to reduce labor and production costs of livestock farms and improve productivity. This technology can be used for rapid pregnancy diagnosis of sows; the location and size of the gestation sacs of sows are directly related to the productivity of the farm. In this study, a system proposes to determine the number of gestation sacs of sows from ultrasound images. The system used the YOLOv7-E6E model, changing the activation function from SiLU to a multi-activation function (SiLU + Mish). Also, the upsampling method was modified from nearest to bicubic to improve performance. The model trained with the original model using the original data achieved mean average precision of 86.3%. When the proposed multi-activation function, upsampling, and AutoAugment were applied, the performance improved by 0.3%, 0.9%, and 0.9%, respectively. When all three proposed methods were simultaneously applied, a significant performance improvement of 3.5% to 89.8% was achieved.

Keywords: Deep learning, Object-detection algorithm, Pig Sac, Sow, Ultrasound

Introduction

As population and income levels continue to rise, there is a corresponding increase in meat consumption. From 2000 to 2019, the per capita consumption of meat increased by 22.7 kg and an average of 2.96% annually in Korea [1]. However, during the same period, the number of farms decreased by 376, and the number of farmers decreased by 1,786,000. The aging population in the farm, with an increase of 24.9% in individuals aged 65 or older (from 21.7% to 46.6% [2]), has contributed to the decline in the labor force. The resultant decline in labor force led to a 13.3% decline in meat sufficiency rate, from 78.8% to 65.5% [1]. In response, the intelligent livestock industry began to incorporate Information and Communications Technology (ICT) in 2014. ICT helps reduce production costs and labor requirements and improve the productivity of livestock farmers. As shown in Figure 1, the number of intelligent livestock farms, which include pig farms, was 23 in 2014 and increased rapidly to 1,073 in 2019 [3]. Equipment such as temperature sensors, humidity sensors, weight scales, and feed management systems are used in the pig farms for pregnant sow management. However, to use these devices optimally, it is necessary to diagnose pregnancy as soon as possible.

There are various methods for diagnosing pregnancy in sows, one of which is measuring urinary and plasma estrone sulfate concentration [4]. This study aims to diagnose pregnancy on sow by analyzing estrone sulfate concentration in plasma and urine. Estrone sulfate concentration in urine was corrected for dilution by creatinine concentration and specific gravity. High performance was achieved in diagnosing pregnancy through estrone sulfate concentration in plasma and urine. Pregnancy diagnosis in plasma and urine recorded recall values of 98.8% and 96.4%, respectively. A study investigated the concentrations of progesterone, estrone, and oestradiol-17β during pregnancy and parturition in sows [5]. When sows were pregnant, progesterone concentrations initially increased and then stabilized. In the case of estrone, it rose during the early and middle stages of pregnancy and decreased just before farrowing, while oestradiol-17β decreased during the early and middle stages of pregnancy and then increased immediately before delivery. Pregnancy was also diagnosed using ultrasound [6]. Unlike other methods, ultrasound pregnancy diagnosis is non-invasive and can minimize stress in sows. Ultrasound images are also mainly used in fetal head and brain analysis [7-8]. A relatively accurate diagnosis is possible even 20 days after mating. Early pregnancy diagnosis is beneficial to the farm, as miscarriages in sows can be reduced by providing necessary nutrition to sows in time [9]. Sow pregnancy must be detected for proper feeding management or antibiotic control to be implemented. Failure to detect sow pregnancy in time increases the non-productive days of sows and causes significant damage to farms [10].

Estimating the number of gestational sacs in pregnant sows is also important when diagnosing pregnancy through ultrasound imaging. The number of gestational sacs can predict litter size and piglet size in a sow, and when combined with the sow's parity number and age, this information offers valuable insights for farm management [11-12]. Based on these studies, an artificial intelligence system is proposed to detect the number and location of gestational sacs in ultrasound images of pregnant sows. This system can provide additional helpful information to pig farmers by identifying the number and location of gestational sacs in pregnant sows. This system is based on an object-detection-based model, whose accuracy was improved through various experiments based on the YOLOv7-E6E model [13]. First, the upsampling technique used in YOLOv7-E6E was modified, and the activation function in the middle of each model was altered. In addition, a data augmentation method was used to increase the amount of data.

Materials and Methods

Dataset

Trained experts collected sow ultrasound data from the National Institute of Animal Science (NIAS) in Cheonan. This study was approved by the Institutional Animal Care and Use Committee (IACUC) of Rural Development Administration (approval No. NIAS-2021-538). Data were collected with MyLab™ OmegaVET (Esaote), and an AC2541 (Esaote) probe with in a frequency range of 1.0Mhz to 8.0Mhz was used. Data were collected in the GEN-M (4.0Mhz-6.0Mhz range frequency) format, often used in pig farms. Data collected by experts between days 23 and 28 post-mating from 103 gestational sows with visible gestational sacs were collected by experts. 4,143 lossless and uncompressed BMP format images were extracted to minimize data loss. Trained experts verified the extracted images and annotated the location of the gestational sacs in each image as bounding boxes.

The 4,143 images were divided into training, validation, and testing sets by randomly splitting them using an approximately standard 6:2:2 ratio, ensuring no data duplication in each dataset. This resulted in 2,484; 828; and 831 images in training, validation, and testing sets, respectively.

Deep-Learning Object-Detection Algorithm

This study aimed to detect and count gestational sacs in ultrasound images using the YOLOv7-E6E model [11]. The YOLOv7-E6E model is a fast and accurate method that combines location detection and object recognition. The performance of the model improved by applying four techniques. The first was extended efficient layer aggregation networks (E-ELAN) for efficient learning when training deep-network models. E-ELAN controls and constructs the gradient path relatively efficiently through extend, shuffle, and merge operations. The second is the compound scaling method for model scaling. The compound scaling method enables fast processing speed by changing the ratio of the input channel to the output channel, reducing hardware usage. The third is a method that improves accuracy without increasing inference costs. A planned re-parameterized convolution was proposed, which showed that the residual connection reduced the performance when the parameter was in the transformed layer. RepConv without identity connection (RepConvN) was used to solve this problem. RepConvN is the algorithm used in deep supervision. The lead head is in charge of the final output, and the aux head is an algorithm that assists learning. This algorithm dynamically adjusts and use acceptable labels from the lead head and coarse labels from the aux head. The last method is mosaic augmentation. The concept of mosaic is straightforward: it involves merging four images into one. This is achieved by resizing each of the four images, stitching them together, and randomly selecting a cutout from the resulting composite to create the final mosaic image. As a result, the objects in the merged image appear at a smaller scale than the original image. This kind of augmentation is beneficial in improving the detection of small objects in images. Performing the mosaic augmentation with the YOLOv7-E6E algorithm poses a challenge in handling the bounding boxes for the final image. Although resizing and relocating the bounding boxes is a manageable task, it can be tedious to determine the appropriate positioning for the boxes after stitching the images together and creating the cutout. In Figure 2, an image is created by mosaic augmentation, and the bounding box marked is the part where the gestational sac is located. This method enabled stable learning even in with small batch size in batch normalization.

The system focused on the structures used in the backbone and head in the YOLOv7-E6E model. First, the model Applied ReOrg to reshape the initial model and the convolution block in the backbone for preprocessing. Then, the process illustrated by the structure in Figure 3 (a) was repeated five times. In the head, after passing through the SPPCSPC layer in which SPP (Spatial Pyramid Pooling) and CSP (Cross-Stage Partial connections) are combined, the processes illustrated by the structures in Figures 3 (b), 3(c), and 3 (d) were repeated three, three, and two times, respectively. Finally, IAuxDetect, which detects object layers, was used. In Figure 3, Conv means a convolution block, DownC means a convolution for downsampling, Shortcut means a layer for residual connection, and Concat means a layer that concatenates multiple feature maps created through convolution.

Multi-Activation Function Method

The activation function is used to transform the model input into output, and a non-linear function is mainly used. The activation function can alleviate the vanishing gradient problem in the deep-learning models, and model configuration can be relatively complex [14]. There are various activation functions, and SiLU, SeLU, ELU, Leaky_ReLU, Mish, and ReLU were used in this study [15-20]. Yolov7-E6E which is used in this study has an activation function in the convolution block, and SiLU was used.

The system combined several activation functions to increase performance in the object-detection model. Iandola et al. [21] improved accuracy and speed using ReLU and PReLU activation functions. Wu et al. [22] improved accuracy and speed using a combination of ReLU and Leaky ReLU activation functions. Based on previous studies, the system proposed the following method.

SiLU and Mish are nonlinear activation functions that add nonlinearity to the neural network. There is a big difference in that SiLU is defined as sigmoid, and Mish is defined as tanh. These differences lead to differences in convergence speed and computational complexity. In general, Mish has a faster convergence speed and higher computational complexity. So, the activation function at the back of the convolution block repeated in the backbone and head of the YOLOv7-E6E model was replaced by Mish. The backbone was modified as shown in Figure 4 (a), and the head was modified as shown in Figure 4 (b), (c), (d) to improve performance.

Up Sampling Method

In the YOLOv7-E6E model used in this study, upsampling was performed three times at the head. Upsampling is a layer that upsamples feature maps according to a stride multiple. In YOLOv7-E6E, the stride multiple is fixed at two; the width and height are doubled through this layer. Upsampling techniques include nearest, bilinear, and bicubic. Nearest is a method of copying the value of the nearest-neighbor pixel. Bilinear is a method of calculating values by performing linear interpolation on each of the two axes using four neighboring pixel values, whereas bicubic calculates a value using a 3rd-order polynomial as an interpolation function using 16 neighboring pixel values [23].

In the YOLOv7-E6E model, the nearest technique was used for all three upsampling. However, nearest is a method of simply copying values; thus, detailed information on the feature map may be lost. Therefore, in this study, the performance was improved by applying a bicubic technique, which has a slightly high computational cost but has low loss and can improve the quality of the feature map.

Augmentation Method

In this study, data were augmented using Google's AutoAugment augmentation technique to improve model performance using a small amount of data [24]. AutoAugment is a reinforcement learning algorithm that automatically searches for improved data augmentation policies. It applies several augmentation techniques in pairs. When a model is trained by applying various augmentation techniques on CIFAR-10, ImageNet, and SVHN datasets, 25 pairs of combinations with the highest performance are disclosed [25-27]. There are 16 augmentation techniques used in AutoAugment: Cutout and Sample Pairing augmentation techniques and Rotate / Shear X, Y / Translate X, Y to rotate, twist, or move the image; Auto Contrast, Invert, Equalize, Solarize, Posterize, Contrast, Color, Brightness, and Sharpness techniques that adjust the image contrast and brightness while the position is fixed.

The CIFAR-10 dataset consists of 32x32 images and is a public dataset with ten classes (Cat, Dog, Frog, Horse, Airplane, Ship, Deer, Bird, Car, and Truck). The ImageNet dataset consists of 1,000 classes of images of various sizes, and the SHVN dataset is a numerical dataset collected from Google street view.

In this study, images were augmented according to the ImageNet augmentation policy. The ImageNet augmentation policy was tuned to a large and diverse dataset. Therefore, unlike the CIFAR-10 or the SVHN augmentation policy, the ImageNet augmentation policy is well generalized. Therefore, the ImageNet augmentation policy expects to perform well in gestational sac detection. The augmented images are shown in Figure 5. They were multiplied 25 times the original amount. The number of images in the training set increased from 2,484 to 62,000, and that of the validation set increased from 828 to 20,700.

A deep-learning model was proposed to detect the gestational sac from ultrasound images of pregnant sows. Three methods were applied to improve its performance. The flowchart of our system is shown in Figure 6.

Results and Discussion

Evaluation Metrics

In this study, mean average precision (mAP) was used as an indicator for comparing the performance of deep-learning models. It is an evaluation index mainly used in deep-learning object detection and measures the similarity between the objects predicted by the object-detection model and the actual object; thus, mAP evaluates the accuracy of the object-detection model. This metric calculates the precision-recall (PR)-curve using precision and recall and the PR-curve to obtain the average precision (AP). AP is calculated as the area under the PR-curve. The mAP can be obtained through the average AP of the class [28]. The model evaluated based on intersection over union (IoU) 0.5. Therefore, only bounding boxes with IoU values greater than 0.5 were calculated.

Multi-Activation Function Result

First, the performance with various activation functions was compared. When the activation function of the convolution block was SiLU, the mAP was 86.3%. When SeLU, ELU, Leaky_ReLU, Mish, and ReLU were consecutively applied, mAP results of 78.1%, 85.7%, 85.6%, 86.0%, and 85.6%, respectively, were achieved the performance evaluation results based on activation functions are summarized in Table 1. SiLU achieved the best result, followed by Mish. The two activation functions of the previously proposed multi-activation function were selected as SiLU and Mish. When the two activation functions were applied, a mAP of 86.6% was achieved, 0.3% more than that of SiLU alone.

Up Sampling Result

Following are the results of comparing upsampling techniques. When nearest was used as the three upsampling techniques at the head of the original model, mAP was 86.3%. When bilinear and bicubic interpolation methods were applied, mAP was 86.5% and 86.5%, respectively, an improvement of 0.2% from the original. The two methods that showed better performance were reconfirmed by applying the previously proposed multi-activation function technique. The mAP of bilinear and bicubic under the multi-activation function application was 86.6% and 87.2%, respectively, improvements of 0.3% and 0.9% from the original model. The results of the evaluation of upsampling methods are presented in Table 2.

AutoAugment Result

Finally, the results present learning and testing augmented images using AutoAugment. Learning and testing using the original data achieved mAP of 86.3% whereas training and testing the model using AutoAugment's ImageNet augmentation policy improved performance by 0.9% to 87.2%. Additionally, Cifar-10 augmentation policy was applied, and it also improved performance by 0.2% to 86.5%. However, ImageNet augmentation policy is better than Cifar-10. The evaluation results are summarized in Table 3. The results showed a significant performance improvement compared to other techniques. More than the original data was needed to train the deep-learning model. The performance was significantly improved because it was trained with a 25 times larger dataset than the original data through augmentation.

Proposed Method Result

When all three methods mentioned above were applied, a mAP of 89.8% was achieved, showing a performance improvement of 3.5% from the original result, which was 86.3%. Each method improved the performance by not more than 1.0%, but the improvement was significant when the three methods were combined. The overall performance of the proposed method is shown in Table 4.

The YOLOv7-E6E-based algorithm used in this study showed high performance in gestational sac detection. First, by modifying the activation function to the multi-activation function, the original model expressed more complex patterns when updating the weights. In addition, when overfitting occurs with one activation function in a specific situation, it can be solved by using another activation function. Therefore, the performance is better than that of the original model. Next, the performance was improved by modifying the upsampling method. It was confirmed that bicubic extracts feature maps with less loss and better quality than bilinear and nearest and improved performance when extracting feature maps. The best performance was obtained by combining all three performance improvement methods. In this study, it is demonstrated that the fusion of the three technologies above has a synergistic effect, significantly improving the model's overall performance. A Multi-Activation function strategy incorporating multiple complex activation functions facilitates broadening the model's nonlinearity. Nevertheless, it is easy to overfit the model due to the complexity of the underlying equations and changes in the parameters. As a result, this tends to bias the learning process toward the training data, even without proper training. However, overfitting can effectively be reduced by the upsampling method and data augmentation techniques. This results in a more robust and accurate model being generated.

The mAP is an index that confirms how similar precisely the model predicts the size and location of the bounding box. As mentioned above, the litter size and the size of piglets can be predicted through the size and position of the gestational sac in the ultrasound image [11-12]. Thus, the improvement in the mAP performance of the model proposed in this study is of great significance. In addition, it is expected to improve the productivity of farms by providing meaningful information to farms.

Conclusion

This study aimed to detect the gestational sac in ultrasound images of sows. Ultrasound images of sows were collected and annotated by experts. A YOLOv7-E6E model was modified by multi-activation function and upsampling methods and trained using this dataset. AutoAugment's ImageNet augmentation policy is used for small amounts of data to improve the deep-learning model's performance. Multi-activation function, changed upsampling method, and image augmentation showed performance improvements of 0.3%, 0.9%, and 0.9%, respectively. When all three methods proposed in this study were applied, there was a significant performance improvement of 3.5%.

In future research, planning to apply a method further to increase performance is necessary. When an image is augmented, there is a case in which the characteristics of the object are not reflected in the augmented image. Therefore, the augmented image may need to be filtered. To improve the performance by filtering out unsuitable augmented images, which do not reflect the characteristics of the object. In addition, the ultrasonic device used in this study is a high-end device manufactured for research purposes, and not a device typically used in farms. However, collecting data with high-end devices is costly and impractical. Data collected with devices commonly used by farmers may add harsh noise and reduce the clarity of the image. Therefore, to solve this problem, additional data collected from devices with low specifications are needed; alternatively, noise generated from devices with low specifications may be added.

Acknowledgments

References

1. Jeong, M. K., Kim, H. J., & Lee, H. W. (2020). Consumer behavior for meat consumption and tasks to respond to its changes. Korea Rural Economic Institute, Naju, Korea.

2. Korean Statistics Information Service. (2022). 『Agricultural Survey』, 『Agricultural Census』

3. Ministry of Agriculture, Food and Rural Affairs, "Current Status of Smart Farm (2021)" https://www.mafra.go.kr/home/5281/subview.do

4. Atkinson, S., & Williamson, P. (1987). Measurement of urinary and plasma estrone sulphate concentrations from pregnant sows. Domestic animal endocrinology, 4(2), 133-138.

5. Cunningham, N. F., Hattersley, J. J., & Wrathall, A. E. (1983). Pregnancy diagnosis in sows based on serum oestrone sulphate concentration. The Veterinary record, 113(11), 229-233.

6. Williams, S. I., Piñeyro, P., & de la Sota, R. L. (2008). Accuracy of pregnancy diagnosis in swine by ultrasonography. The Canadian Veterinary Journal, 49(3), 269.

7. Torres, H. R., Morais, P., Oliveira, B., Birdir, C., Rüdiger, M., Fonseca, J. C., & Vilaça, J. L. (2022). A review of image processing methods for fetal head and brain analysis in ultrasound images. Computer Methods and Programs in Biomedicine, 106629.

8. Alzubaidi, M., Agus, M., Shah, U., Makhlouf, M., Alyafei, K., & Househ, M. (2022). Ensemble Transfer Learning for Fetal Head Analysis: From Segmentation to Gestational Age and Weight Prediction. Diagnostics, 12(9), 2229.

9. Einarsson, S., Madej, A., & Tsuma, V. (1996). The influence of stress on early pregnancy in the pig. Animal Reproduction Science, 42(1-4), 165-172.

10. Koketsu, Y., Tani, S., & Iida, R. (2017). Factors for improving reproductive performance of sows and herd productivity in commercial breeding herds. Porcine health management, 3(1), 1-10.

11. Kousenidis, K., Giantsis, I. A., Karageorgiou, E., & Avdi, M. (2021). Swine ultrasonography numerical modeling for pregnancy diagnosis and prediction of litter size. Int. J. Biol. Biomed. Eng, 15, 29-35.

12. Kousenidis, K., Kirtsanis, G., Karageorgiou, E., & Tsiokos, D. (2022). Evaluation of a Numerical, Real-Time Ultrasound Imaging Model for the Prediction of Litter Size in Pregnant Sows, with Machine Learning. Animals, 12(15), 1948.

13. Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv preprint arXiv:2207.02696.

14. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

15. Elfwing, S., Uchibe, E., & Doya, K. (2018). Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. Neural Networks, 107, 3-11.

16. Klambauer, G., Unterthiner, T., Mayr, A., & Hochreiter, S. (2017). Self-normalizing neural networks. Advances in neural information processing systems, 30.

17. Clevert, D. A., Unterthiner, T., & Hochreiter, S. (2015). Fast and accurate deep network learning by exponential linear units (elus). arXiv preprint arXiv:1511.07289.

18. Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013). Rectifier nonlinearities improve neural network acoustic models. In Proc. icml (Vol. 30, No. 1, p. 3).

19. Misra, D. (2019). Mish: A self regularized non-monotonic activation function. arXiv preprint arXiv:1908.08681.

20. Glorot, X., Bordes, A., & Bengio, Y. (2011). Deep sparse rectifier neural networks. In Proceedings of the fourteenth international conference on artificial intelligence and statistics (pp. 315-323). JMLR Workshop and Conference Proceedings.

21. Iandola, F., Moskewicz, M., Karayev, S., Girshick, R., Darrell, T., & Keutzer, K. (2014). Densenet: Implementing efficient convnet descriptor pyramids. arXiv preprint arXiv:1404.1869.

22. Wu, B., Iandola, F., Jin, P. H., & Keutzer, K. (2017). Squeezedet: Unified, small, low power fully convolutional neural networks for real-time object detection for autonomous driving. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 129-137).

23. Gonzalez, R. C., & Woods, R. E. (2008). Digital image processing. Prentice Hall.

24. Cubuk, E. D., Zoph, B., Mane, D., Vasudevan, V., & Le, Q. V. (2018). Autoaugment: Learning augmentation policies from data. arXiv preprint arXiv:1805.09501.

25. Krizhevsky, A., & Hinton, G. (2009). Learning multiple layers of features from tiny images.

26. Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255).

27. Netzer, Y., Wang, T., Coates, A., Bissacco, A., Wu, B., & Ng, A. Y. (2011). Reading digits in natural images with unsupervised feature learning.

28. Everingham, M., Van Gool, L., Williams, C. K., Winn, J., & Zisserman, A. (2010). The pascal visual object classes (voc) challenge. International journal of computer vision, 88, 303-338.

Tables and Figures

Table 1. Performance evaluation of activation functions.

Dataset	Activation Function	mAP
Original	SiLU (original)	0.863
	SELU	0.781
	ELU	0.857
	LeakyReLU	0.856
	Mish	0.860
	ReLU	0.856
	Proposed Method (SiLU + Mish)	**0.866**

Table 2. Performance evaluation of upsampling methods.

Dataset	Activation Function	Upsampling	mAP
Original	SiLU (original)	nearest (original)	0.863
	SiLU (original)	bilinear	0.865
	SiLU (original)	bicubic	0.865
	Proposed Method (SiLU + Mish)	bilinear	0.866
	Proposed Method (SiLU + Mish)	**bicubic**	**0.872**

Table 3. Performance evaluation of augment method.

Dataset	Upsampling	mAP
Original	nearest	0.863
Cifar-10 Augmentation	**nearest**	**0.865**
ImageNet Augmentation	**nearest**	**0.872**

Table 4. Overall performance evaluation of proposed method.

Dataset	Activation Function	Upsampling	mAP
Original	SiLU (original)	nearest	0.863
ImageNet Augmentation	**Proposed Method (SiLU + Mish)**	**bicubic**	**0.898**

Figure 1. Number of intelligent livestock farms by year.

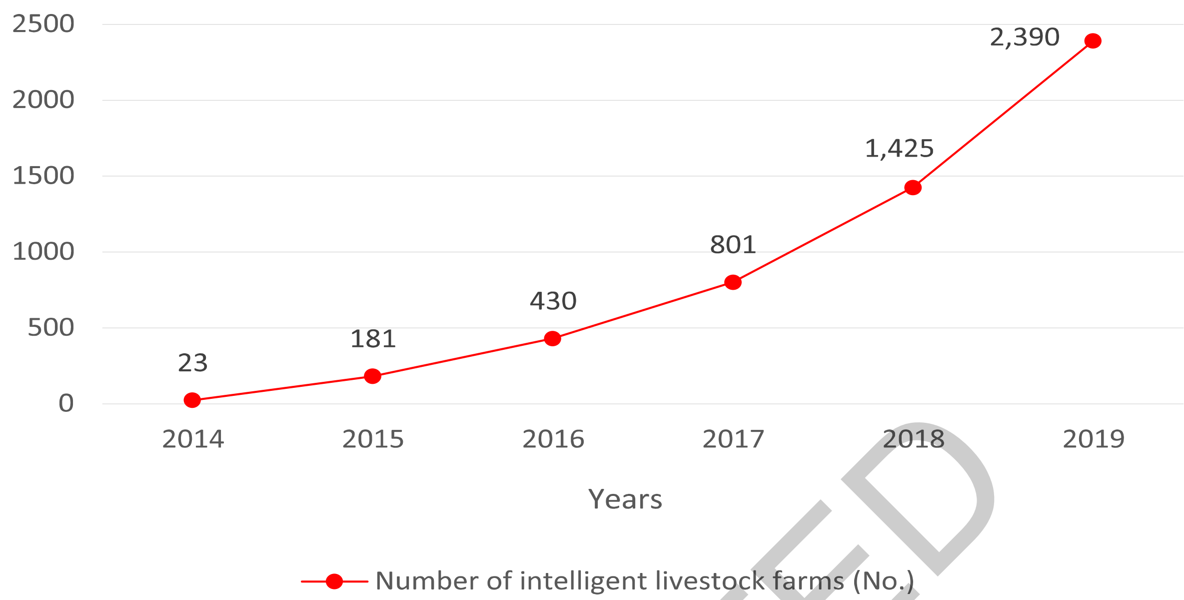

Figure 2. Example of mosaic augment.

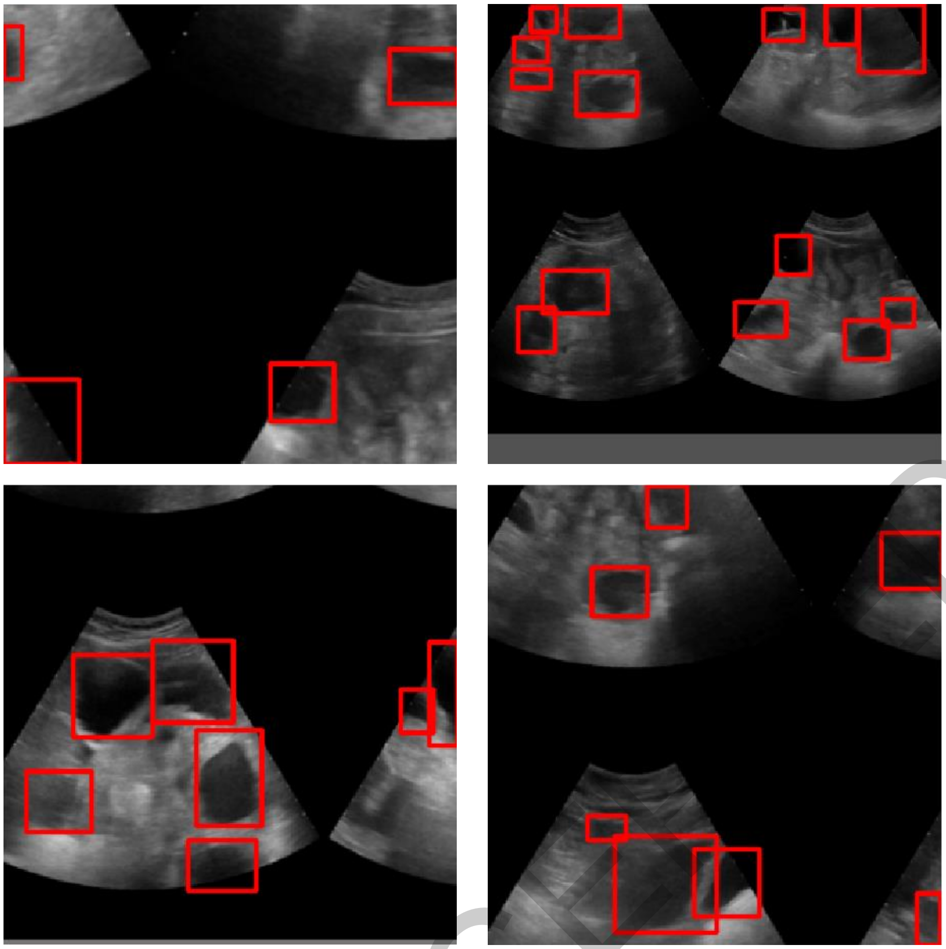

Figure 3. Original model (YOLOv7-E6E).

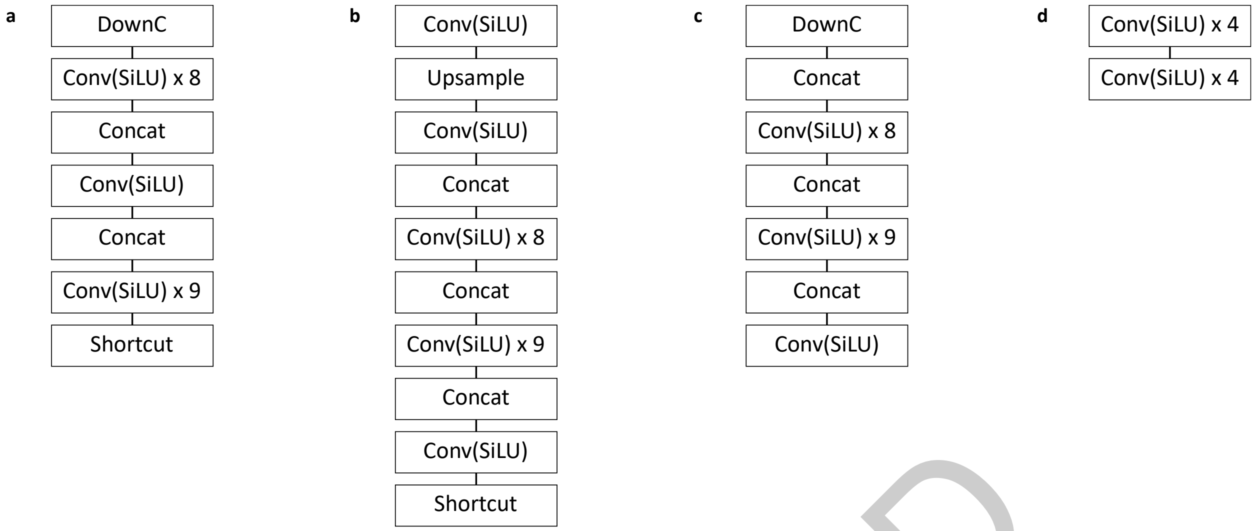

Figure 4. Proposed model (activation function changed).

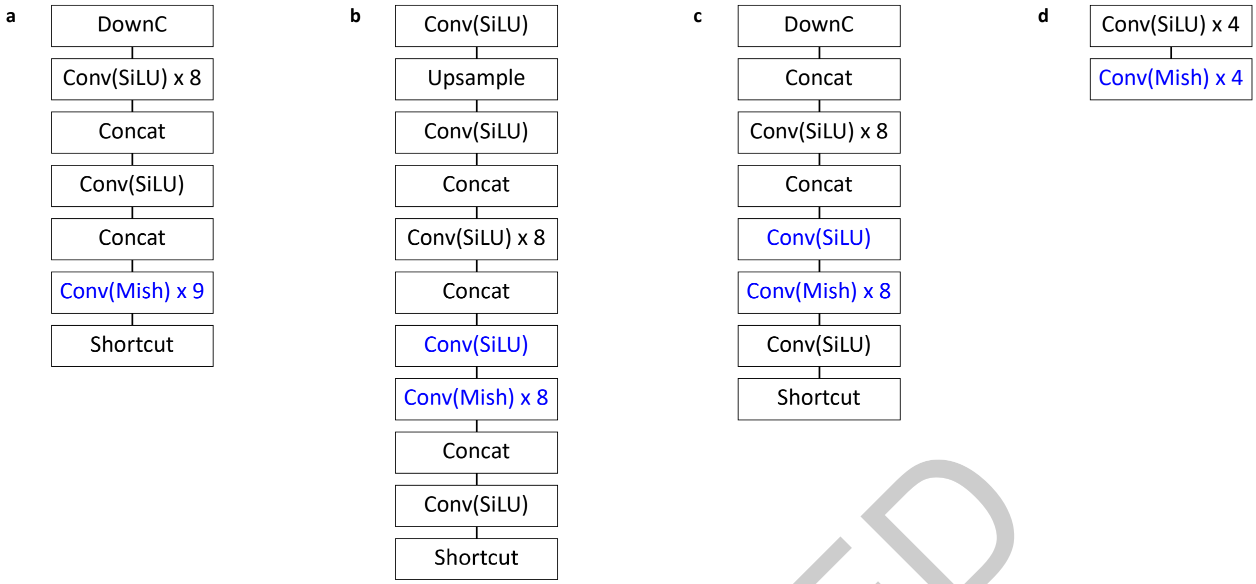

Figure 5. Example of augmented data.

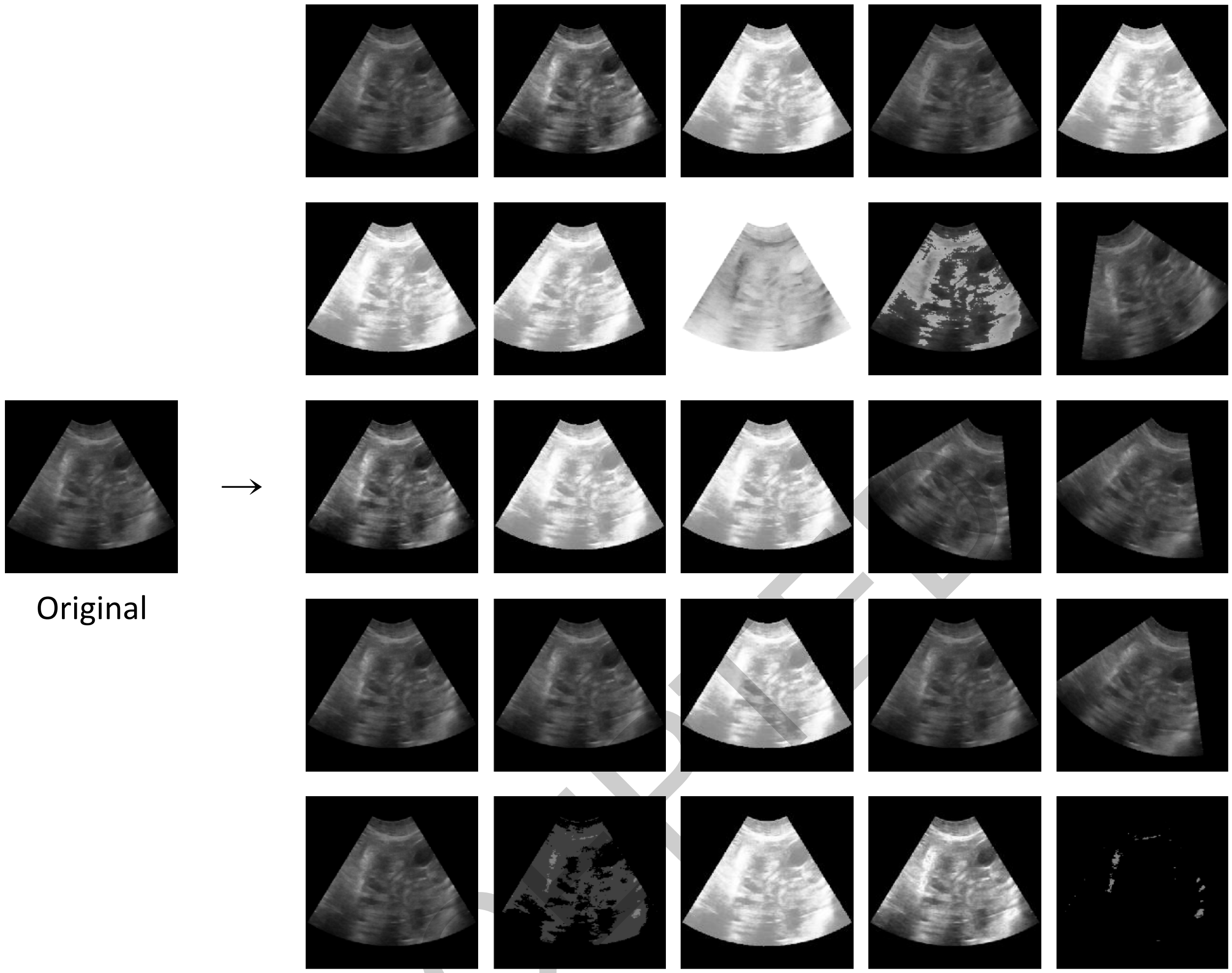

Figure 6. Flowchart of the proposed scheme.